PLAN FOR IMPROVING THE QUALITY OF A NAVAL COMBAT SYSTEM TEST AND INTEGRATION PROCESS

by

John M. Stroud

Report submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

SYSTEMS ENGINEERING

APPROVED:

Konstantinos P. Triantis
K. P. Triantis, Chairman

Ben S Blanchard
B. Blanchard

W. Canaday

May, 1991

Blacksburg, Virginia

c.2

PLAN FOR IMPROVING THE QUALITY OF A NAVAL COMBAT SYSTEM TEST AND INTEGRATION PROCESS

by

John M. Stroud

Committee Chairman: Dr. K. P. Triantis
Industrial and Systems Engineering

(ABSTRACT)

A naval Combat System comprises several subsystems or elements. The Combat System test and integration process ensures the elements function properly as a system. Naval Combat Systems are complex, costly, schedule driven, and have many independent activities that contribute to their development. Because of these characteristics, the system test and integration process is formidable.

A typical Combat System test and integration process is outlined. Historical examples of naval Combat System integration efforts are given. Obstacles to the system test and integration process are described and solutions for the obstacles are suggested. In order to measure the improvement, or deterioration, of the test and integration process, several metrics that quantify the quality of the process will be discussed.

ACKNOWLEDGEMENTS

I would like to thank my employer, the Naval Surface Warfare Center, for giving me the opportunity to continue my education.

I would especially like to thank Bill Canaday, whose insight and experience provided me with the guidance to make this a better report.

I must also thank Dr. Konstantinos Triantis, who showed me the importance of quality. His commitment to his students is a tribute to the University.

TABLE OF CONTENTS

Chapter

Chapter

LIST OF ILLUSTRATIONS

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

ACSC	AEGIS Combat System Center
ASU	Approved for Service Use
C&C	Command and Control
CDR	Critical Design Review
CIC	Combat Information Center
CM/QA	Configuration Management/Quality Assurance
COMOPTEVFOR	Commander, Operational Test and Evaluation Force
CNO	Chief of Naval Operations
CPCR	Computer Program Change Request
CPDD	Computer Program Definition Document
CSED	Combat System Engineering Development
CSLBTS	Combat System Land Based Test Site
CSM	Combat System Manager
CSTP	Combat System Test Plan
DOD	Department of Defense
FCDSSA	Fleet Combat Direction Systems Support Activity
IDS	Interface Design Specification
ITP	Integrated Test Package
LBTS	Land Based Test Site
LSE	Lifetime Support Engineering

LSEA	Lifetime Support Engineering Agent
MOE	Measures of Effectiveness
NGT	Nominal Group Technique
NOSC	Naval Ocean Systems Center
NSWC	Naval Surface Warfare Center
OIC	Officer in Charge
PARM	Participating Manager
PCO	Prospective Commanding Officer
PTC	Production Test Center
SHAPM	Ship Acquisition Project Management
STDR	System Test Disclosure Review
ST&I	System Test and Integration
T&E	Test and Evaluation
TEMP	Test and Evaluation Master Plan
TQM	Total Quality Management
WCS	Weapons Control System
WFA	Work Flow Analysis

CHAPTER 1

INTRODUCTION

Introduction

A system can be defined as "an assemblage or combination of elements or parts forming a complex or unitary whole."[1] The system under consideration here is actually a process. The process is U.S. Naval Combat System Test and Integration (ST&I). Integration testing is an orderly progression of testing in which software and/or hardware elements are combined and tested until the entire system has been integrated.[2] The purpose of ST&I is to conduct risk assessment for the integrated system so management can make the right decisions. These decisions include whether or not the system is ready for delivery to the fleet.

The Combat System for a specific platform includes the collection of those sensor, command and control, and weapon subsystems installed in that platform.[3] The Combat System includes both the hardware and all associated software. Software is the most volatile component of the Combat System. It constantly undergoes change to meet new threats, incorporate technological advances, correct problems, and adapt to new hardware. Because of this constant change, the test and integration of the software components of subsystems will be the focus of this study.

This study will explore the Lifetime Support Engineering (LSE) phase of ST&I; much of the material may be applied to other phases of system development and evaluation. The LSE phase in the life of a Combat System poses special problems for ST&I that do not exist in the development phase. The LSE phase of a system comes after the system has been designed, developed, undergone extensive development and operational testing, and is operationally deployed. The LSE organization, which may be different than the development organization, is responsible for the maintenance of the Combat System over the remainder of its life.

The LSE organization inherits the design and coding techniques of the system developer. The developer spent a considerable amount of time designing, coding, and testing the system. Because of this, the developer is initially more familiar with the details of the system design than the LSE organization. Good documentation aids the LSE process but it is no substitute for knowledge of the system developers.

Computer programs resulting from the system development phase are installed on ships under construction. The programs undergo extensive test and evaluation at the shipyard and during sea trials before they are ever operational. The LSE organization installs programs on a ship that must be war-ready when the installation is complete.

Capers Jones defines quality software as "software that combines the characteristics of low defect rates and high user satisfaction."[4] Another definition that is often used is "software that conforms to requirements." This latter definition does not take into account incomplete or incorrect requirements -- a common source of software defects. Ensuring a high degree of usability should always be the focus of a Combat System integration process.

The reason to strive for continuous improvement of the ST&I process is to improve the quality of the product. The hypothesis is that the quality of the process and the quality of the product are positively correlated. If the process of ST&I is improved then the quality of risk assessment is improved. All the right tests needed to ensure Combat System integration will be completed within schedule. With more effective testing, decisions with regard to the Combat System can be made with confidence. The better the ST&I process, the less chance a serious problem will reach the fleet.

Objectives

This study will identify those elements that are significant for improving the quality of the Combat System

Test and Integration process and hence improving the quality of the Combat System itself. This process of identifying the "vital few and trivial many" is known as Pareto analysis.[5] By focusing on the major contributors to the improvement of quality, resources are optimized to gain the greatest benefit.

Measurement provides the means for analyzing the performance of an organization or the quality of a product. A metric is a standard of measurement. Metrics must be easily understood, undisputed, and have a feasible means of gathering data. If a metric fails on any one of these points, its usefulness fades. Metrics will be developed to quantify the performance of the ST&I process.

Methodology

Several steps must be taken to complete an analysis of the ST&I process. Chapter 2 analyzes the environment that influences the process. The chapter identifies those factors that make it necessary for an ST&I organization to strive for continuous improvement. These factors are uncontrollable elements in the environments of the Department of Defense, the U.S. Navy, many subsystem development organizations, and current technology.

Chapter 3 describes a typical U.S. Navy Combat System test and integration process. The process is modeled after one used for the test and integration of a current high-technology Combat System. Problems inherent in the process are identified. The Combat System itself is described to aid in the visualization of the integration process.

The analysis would not be complete without examining historic examples of U.S. Navy Combat System test and integration. Chapter 4 outlines lessons learned from past ST&I projects. The study of history helps to prevent the repetition of past mistakes.

Chapter 5 recommends solutions for the problems encountered in Chapters 3 and 4. It is a plan for improving the quality of the ST&I process. Methods that have been successful in other organizations and disciplines are applied to the ST&I process.

Chapter 6 is perhaps the most important of all. It provides methods for measuring the performance of the ST&I process. Without these measures there is no way of knowing whether programs to improve quality have been effective.

A short summary and conclusions from the study are presented in Chapter 7. Avenues for future investigation are discussed.

CHAPTER 2

THE ENVIRONMENT

The ST&I process is influenced by several factors that cannot be controlled within the ST&I organization. These environmental factors must be identified in order to understand their impact on the ST&I process. They make the need for improvement within the organization necessary. By identifying the uncontrollable elements the organization can focus on improving the controllable elements.

Combat System Complexity

The following excerpt is from an article that appeared in the NAVSEA Journal:

> On the afternoon of 23 September 1779, Commodore John Paul Jones made visual contact with a merchant fleet convoyed by a sloop of war and a heavily armed frigate. It was early evening when Commodore Jones ordered the firing of a starboard broadside from his flagship, the Bonhomme Richard, at the 50 gun British frigate. Serapis returned fire and, before long, Jones decided his only chance was to grapple and board the enemy ship. Sailing through a storm of cannonfire, Bonhomme Richard caught Serapis with grappling irons in a bow to stern embrace, muzzles of ships' guns touching. Ships' crews were alternately acting as firemen with buckets and fighting men with muskets and grenades. For over two hours the battle raged until, at last, the British Captain surrendered Serapis to Commodore Jones.
>
> The tactical decisions and combatant equipment utilized in Jones' encounter with Serapis would indicate a well-coordinated Combat System aboard Bonhomme Richard. Early detection and identification of the enemy were

> made, combat decisions were communicated effectively and ships' armaments were successfully brought to bear. Certainly Bonhomme Richard has no place in today's surface combatant fleet. Her armament would obviously be no match for current shipboard missile and gun systems. Similarly, Commodore Jones' spyglass would be a poor substitute for the 3-D radars of today. And finally, the communications between combat subsystems would be painfully inadequate to counter today's threats.[6]

This passage illustrates the tremendous advances in naval Combat System technology over the past 200 years. The Combat Systems of today are indeed more technologically advanced and complex than those of even 20 years ago. This growth in system complexity increases at a tremendous rate. The expanding capability of computers, weapons performance, sensor technology, and intelligence gathering means system requirements also expand. A Combat System is not born, it evolves over its life-cycle.

As systems become more complex the need for automated information exchange and decision making increases. This means the number of subsystem interfaces and the amount of information exchanged between subsystems multiplies. More functions are allocated to computer programs and less to human operators. This is because the computer programs are more effective at assimilating large amounts of data in a relatively small amount of time.

The increase in system capability and requirements directly impacts the system test and integration requirements. As the number of interfaces and the amount of data exchanged goes up, so must the number of tests that evaluate those capabilities. Even though a new system configuration is more complex than a previous one, the amount of time allocated to ST&I by the LSE organization remains fixed. The additional requirements, which must be tested in the same amount of time, make it necessary for the ST&I organization to improve the internal process to accomplish its goals.

System complexity is constantly increasing. In order to keep up with increasing demands, an ST&I organization must strive to continuously improve performance. If the demands are not met, the quality of the product suffers. The product is an integrated Combat System whose performance is critical to the warfighting capability of the ship and the safety of those who operate it. If the Combat System does not perform when called upon, the results may be catastrophic.

Number of Organizations Involved

A typical Combat System may contain 20 or more subsystems (see Chapter 3). Each one of these subsystems, which must be integrated into the Combat System, may be developed by a different Navy or contractor organization. The coordination of these activities can be a logistics nightmare.

While under development, the system configuration and design is rigorously controlled from a central program office with a well defined line authority. As the span between completion of system development and operational support grows so does the number of new requirements incorporated into the system and subsystems. This increases the chance that subsystems will not be compatible.

Each subsystem organization often falls under the direction of a different Navy program office. These program offices have their own goals and objectives that may or may not be consistent with those of the program office for a particular ship class. The same subsystem may be installed in several ship classes. While the subsystem development or upgrade schedule may fit the needs of one ship class, it may not fit the needs of two or more at the same time. This is particularly a problem during the LSE phase as individual schedules drift further apart. Because these organizations take direction from their own program offices, it is difficult for the ST&I organization to influence their policies or procedures.

The subsystem development organizations are dispersed over a wide geographic area. As a result, communications are difficult. Face-to-face meetings are always more productive and meaningful than phone calls; the number of organizations and locations often make these meetings impossible. This means the organizations must work harder to communicate schedules, project objectives, and system incompatibilities. If they do not, lack of cooperation and discontent may result.

The problems associated with coordinating these activities will never go away. The ST&I organization has no control over these outside activities. Even so, the organization must seek ways to improve coordination. This will help to prevent unforseen problems. The ST&I organization must improve those things it can control and adapt to those it cannot.

Schedules

The U.S. Navy's AEGIS Combat System is the most technologically advanced shipboard Combat System in the world. The first AEGIS ship was commissioned in 1983. According to current purchasing schedules, 48 AEGIS ships will have been commissioned by 1997.[7] Because of this 15 year span and evolving technology, the Combat System configurations vary from ship to ship. Capabilities exist for ships under construction today that did not exist in 1983. Because of staggered equipment installation schedules, it would be rare to find two ships with identical Combat System configurations. The differences between ships constructed close together will be small; while the differences between ships constructed further apart may be great.

This illustrates the configuration diversity that can exist within a single ship class. The large number of configurations within the ship class increases the amount of ST&I that must be conducted by the Lifetime Support Engineering Agent (LSEA).

Each ship has a pre-scheduled period, or availability, during which systems and equipment may be installed or upgraded. These periods are scheduled roughly every two years. Because the hardware configuration may change, and with the crew available for familiarization, these periods usually include a major computer program configuration change. With roughly half of the ships in a class having availabilities in any one year, the amount of integration testing required can be enormous.

Subsystem computer program upgrades are interspersed among the major configuration changes. The agencies responsible for developing subsystems often operate on cycles for producing version updates or upgrades. Upgrades for a particular subsystem may occur on 2 or 3 year cycles. Subsystems may be revised more often to meet new threats or provide emergency problem fixes.

Because there are numerous subsystems, all operating on different schedules, the ST&I organization is constantly working to integrate the subsystems in a timely manner. One subsystem upgrade may require a separate ST&I effort for each Combat System configuration. The number of ST&I efforts required is a factor of the number of subsystem changes and the number of unique shipboard configurations.

Because subsystems are developed by independent activities, the ST&I organization has little control over their schedules. Many times subsystem developers are also working to meet a particular ship that is scheduled for upgrade. Time allotted for ST&I, prior to upgrade installation, may be overlooked when subsystem development schedules are created. If ST&I time is allotted and the subsystem schedule changes, plans must be altered to accommodate the change.

The availability of a Land Based Test Site (LBTS) also impacts the effectiveness of ST&I. A LBTS is used to emulate an actual shipboard configuration. It combines real Combat System hardware and simulation programs to create an integrated Combat System test environment.

The LBTS is used for program development, crew training, system demonstrations, testing by subsystem developers, and for ST&I. Because the LBTS is heavily used, it has limited availability. ST&I gets only a fraction of the available LBTS time. As a result, the allocated ST&I periods must be used to maximum efficiency. The ST&I organization cannot afford to waste valuable LBTS time.

A LBTS has many inherent deficiencies. Ship motion must be simulated. The LBTS lacks complete hardware configurations such as: missile launchers, guns, full coverage radar, ship superstructure, etc. Because of the deficiencies that exist at a LBTS, Combat System level shipboard testing is often required. Configurations that do not exist at the LBTS must be tested on board ship. The ship's operational demands, rigid schedule, and ship's support required make the scheduling of shipboard testing difficult. Careful test planning must take place well in advance of the scheduled time. If the planned test date falls through, recovery may be difficult.

Much of the ST&I organization's planning schedule depends on uncontrollable outside activities. The organization must be able to constantly adapt to schedule changes and compressions. Long range planning is difficult because of the number of activities involved and the volatility of individual schedules.

Personnel Factors

ST&I organizations within the Department of Defense (DOD) face the same personnel problems that plague other technology oriented activities in the U.S. Government. The widening pay gap between public and private sector employees makes it difficult to hire and retain skilled workers. Low starting salaries hinder efforts to attract the best college graduates. Shrinking defense budgets may result in DOD wide hiring freezes.

System test engineers possess skills from hands-on experience in all aspects of the Combat System. These skills make them highly sought after in other DOD organizations as well as defense contractor organizations. The job of ST&I can be very demanding. Large amounts of travel to remote test sites and ships is required. Work hours are often erratic to meet LBTS and ship schedules.

It takes years of hands on experience to create a skilled system test engineer. All of the above factors must be overcome to hire and retain skilled workers. The limited amount of personnel available to do the very demanding task of ST&I forces the organization to optimize its resources to meet demands.

CHAPTER 3

A TYPICAL Combat System TEST AND INTEGRATION PROCESS

The Anatomy of a Combat System

Figure 1 shows some of the key subsystems that form a U.S. Navy Combat System. This figure is not meant to describe any particular Combat System; it is meant to depict some of the major communication interfaces and configurations that must be considered when conducting ST&I. A particular ship may not contain all of the subsystems or interfaces shown. There may also be other subsystems and additional interaction between subsystems.

The key subsystems within the core of a Combat System are Command and Control (C&C) and Weapons Control System (WCS). The remainder of the Combat System is built around these cornerstones that are unique to their particular ship class. The life-cycle control of these core elements is performed by the same Navy activity that controls the ST&I organization -- the LSEA.

The C&C subsystem interfaces with all of the sensor, display, communications, and weapons subsystems. It receives data, assesses the data, and directs various subsystems to act on the data. It maintains the status of

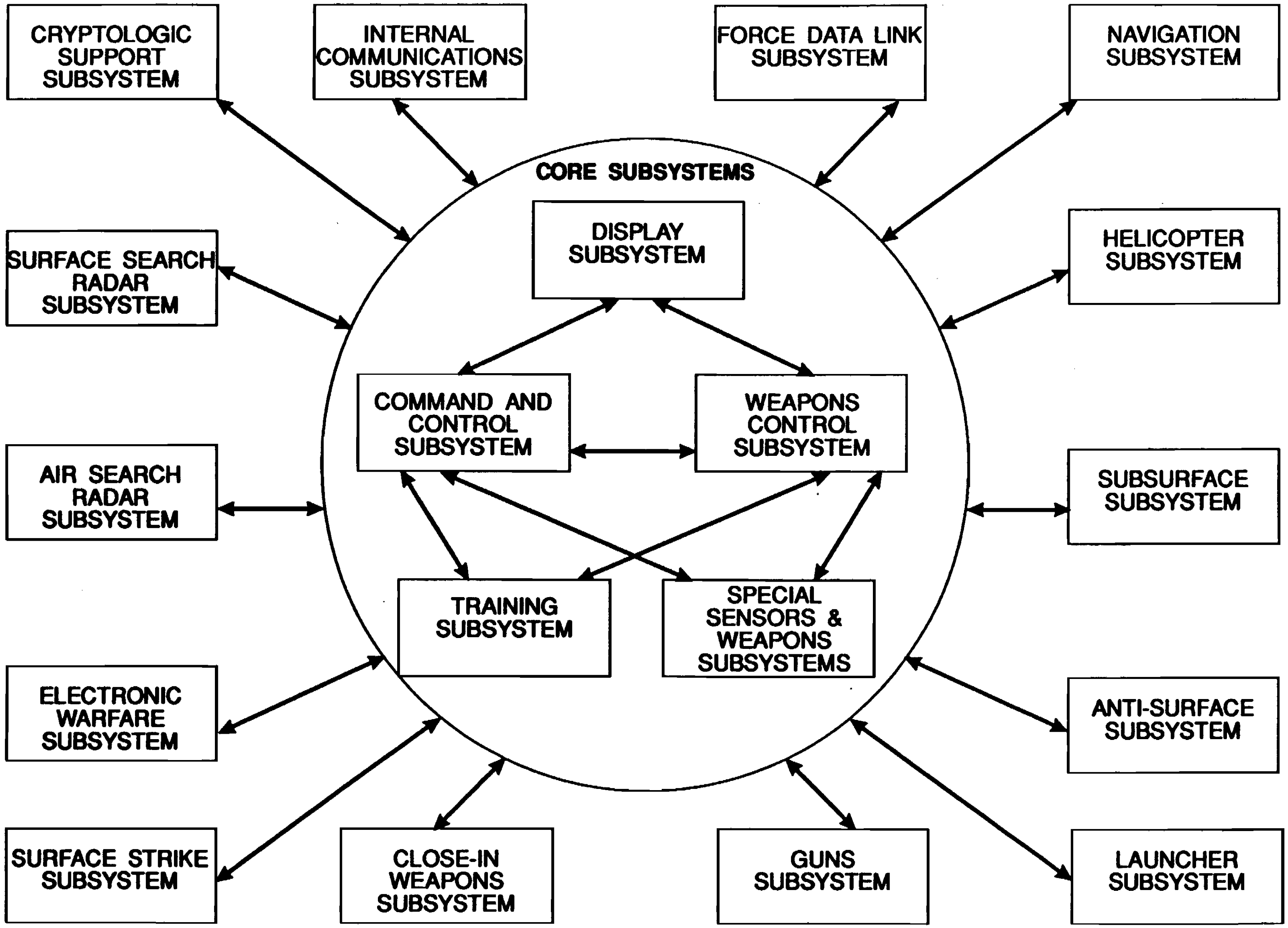

Figure 1. Anatomy of a Combat System

all subsystems. Because the C&C subsystem is the central repository for Combat System information, it is a key interface for many of the subsystems.

The WCS subsystem, which on some ship classes is referred to as the Fire Control System, interfaces with many of the ship's weapons. It receives targeting information from the C&C subsystem and directs the appropriate weapon to the target. The weapons it interfaces with may include: anti-aircraft missiles, close-in weapon subsystems, guns, anti-surface weapons, underwater weapons, and helicopter subsystems. Like C&C, WCS has many subsystem interfaces.

In addition to the C&C and WCS subsystems, there may be additional subsystems that are unique to the ship class and maintained by the LSEA. These other subsystems could include: training programs, data extraction and reduction programs, special sensor or weapon subsystems, and information display programs.

Components of a typical LSEA organization include: individual subsystem development organizations, systems engineering, support functions, Configuration Management and Quality Assurance (CM/QA), and ST&I. These organizations work together to maintain the Combat System. If a change is made within the core elements then a certain amount of testing must be conducted. The extent of testing depends on the degree of impact the change has on the system. Major core upgrades are scheduled so all of the core elements are upgraded at once. This increases Combat System cohesiveness and reduces the number of separate integration events that must be conducted.

The subsystems that surround the core make up the remainder of the Combat System. These subsystems not only interact with the core but may also interact with each other. As stated in chapter 1, these subsystems are maintained by a number of different organizations. There may be secondary subsystems within these subsystems. In general, it is not within the scope of the ST&I organization to test the interaction of these secondary subsystems. Because the interfaces to these secondary subsystems are an interface removed from the core, the testing responsibility lies with the subsystem developer.

The System Test and Integration Process

Figure 2 shows a high-level view of the ST&I process, along with its inputs and outputs, within the environment described in Chapter 1. Included in the figure are: the upstream organizations, the inputs generated from the upstream organizations, the ST&I transformation process, outputs from the transformation process, and the downstream organizations.

The system designers, upstream from the ST&I organization, control the A-specifications and Interface

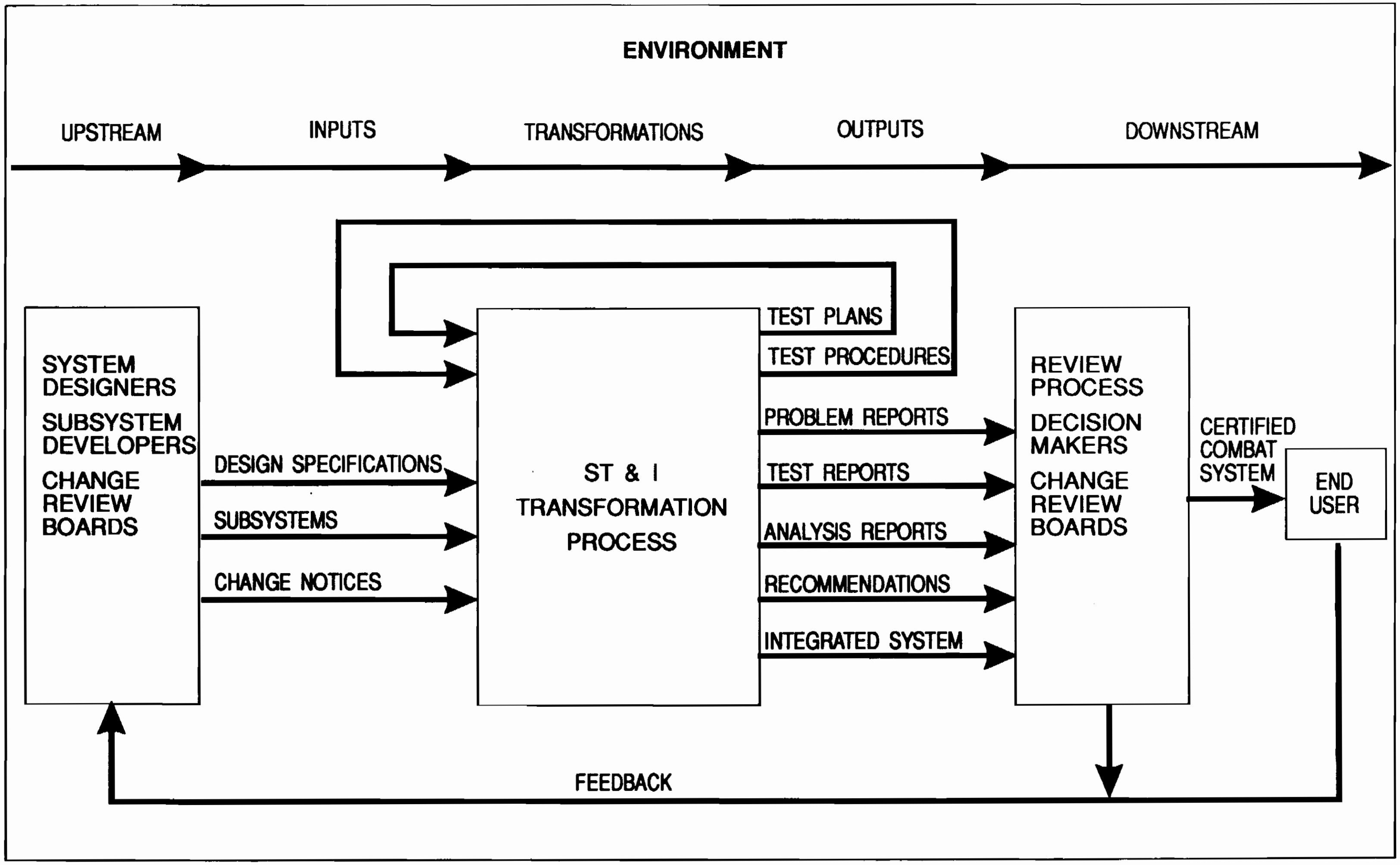

Figure 2. The System Test and Integration Process

Design Specifications (IDS). They coordinate changes between subsystems and plan major functional upgrades or changes. The subsystem developers, both core and non-core, are also upstream from the ST&I process. These organizations are responsible for the development, test, and certification of their respective subsystems. They generate the design specifications (B-specifications), conduct internal change review boards, and provide specification changes for the subsystems. The ST&I organization requires the individual subsystems be certified as operationally ready, by the developer, before the start of ST&I.

The outputs from the upstream organizations are input to the ST&I transformation process. The ST&I organization takes the input, assesses it, develops test plans and procedures, and uses the plans and procedures to generate output. The output is in the form of: test reports, analysis reports, problem reports, an integrated Combat System, and recommendations.

The problem reports are submitted to the change review boards of the core Combat System elements and the non-core elements. The test and analysis reports are submitted to a delivery review board along with the ST&I recommendations. The review board weighs the test report, analysis report, and ST&I recommendations and decides if the Combat System is ready for delivery and installation. If the system is approved for delivery, the ST&I organization is part of a team that installs the computer programs on board ship.

Figure 3 is a more detailed functional flow diagram of the ST&I process. The functional flow diagram is useful here because it shows the flow of the entire process and allows for expansion of those functions that must be shown in greater detail. Functions on the top level, numbered 1.0, 2.0, and 3.0, correspond to the inputs, transformation process, and outputs respectively. If a function is broken down further, the next level of indenture is coded at the next decimal level. For example, functions on the first level of function 2.0 would be 2.1, the next 2.1.1 and so on. Reference blocks are used to show interfacing functions. A reference block is depicted by an open bracket. If a block is a decision block then YES and NO paths will be shown.[8] Figure 3 will be referred to frequently in the following detailed process description.

A major baseline upgrade of the core Combat System elements can be more than a two year effort from conception to delivery. The ST&I organization gets involved when the system level changes have been identified. The ST&I organization receives the A-specification and IDS change packages and considers several factors when developing system level tests for the functional changes (Reference block 2.1). Factors that are considered include: What

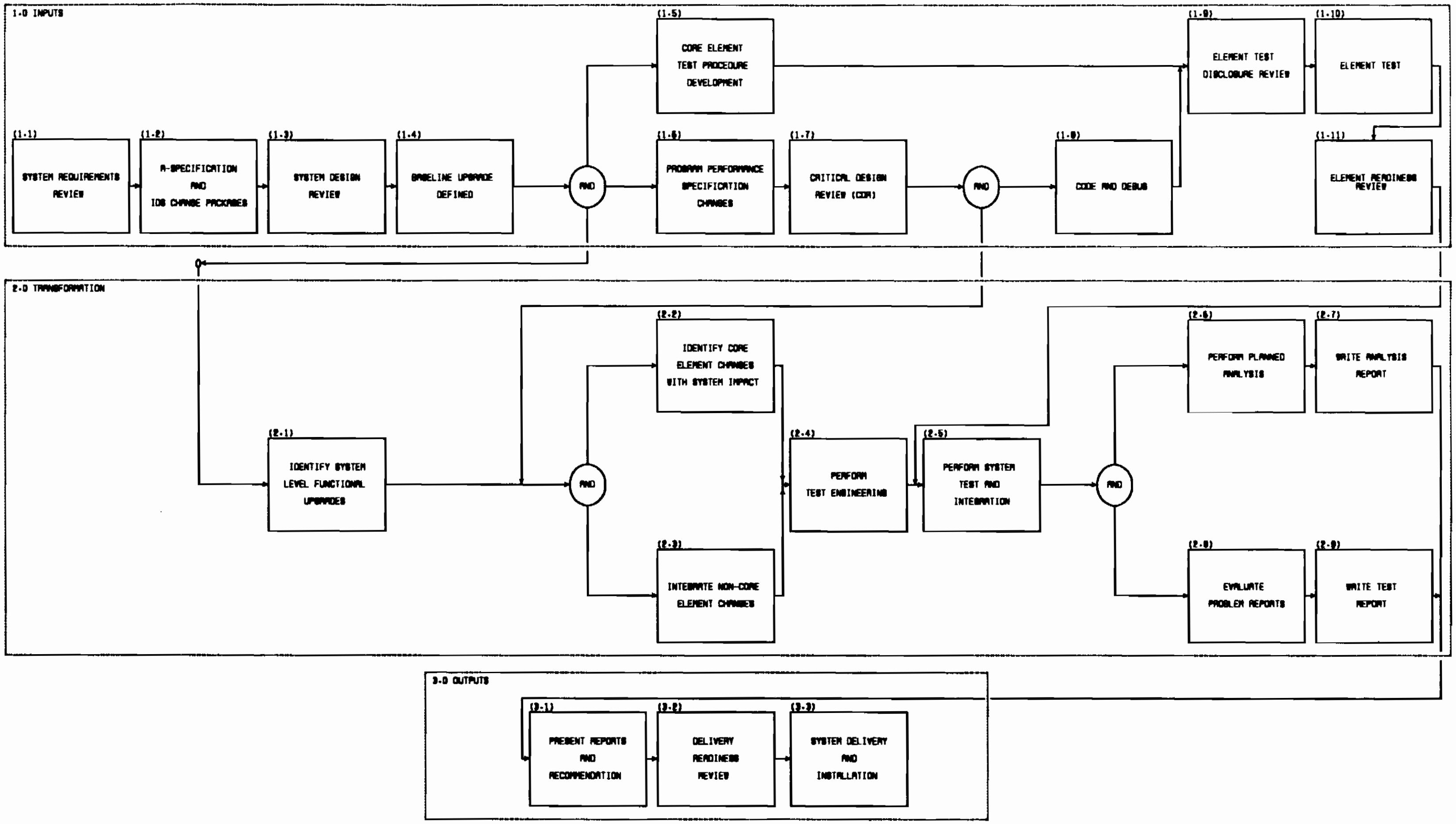

Figure 3. System Test and Integration Functional Flow

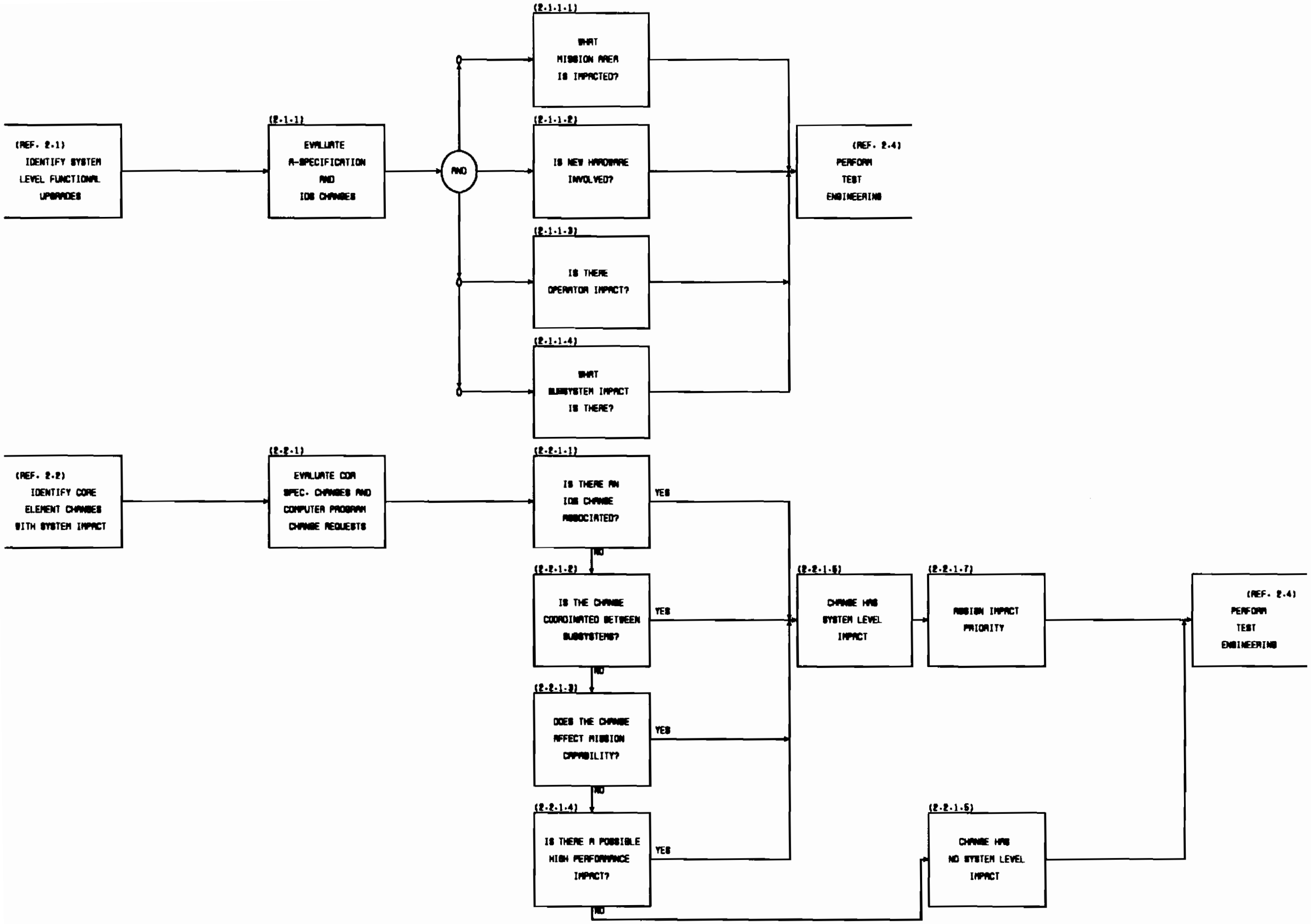

Figure 3. (continued)

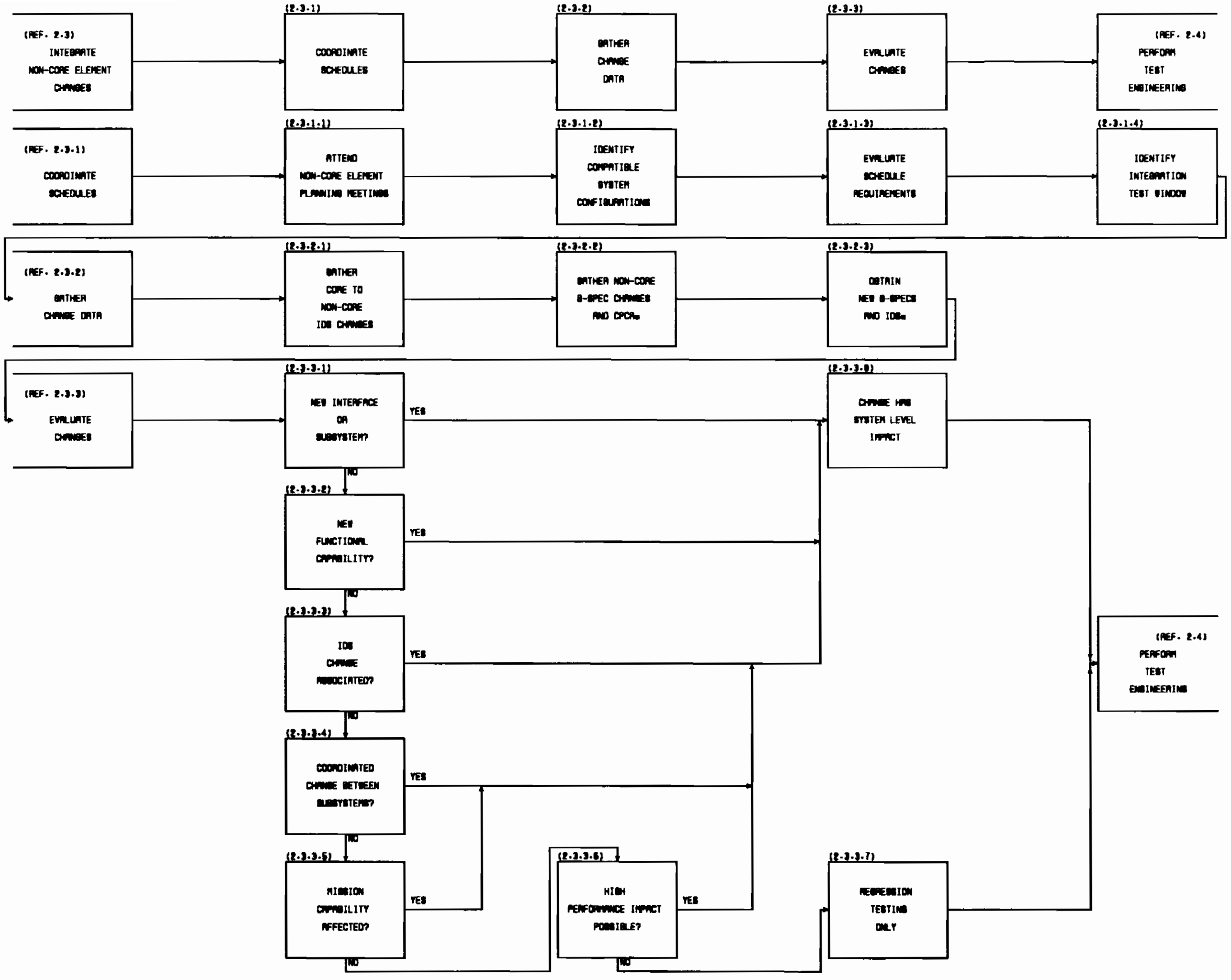

Figure 3. (continued)

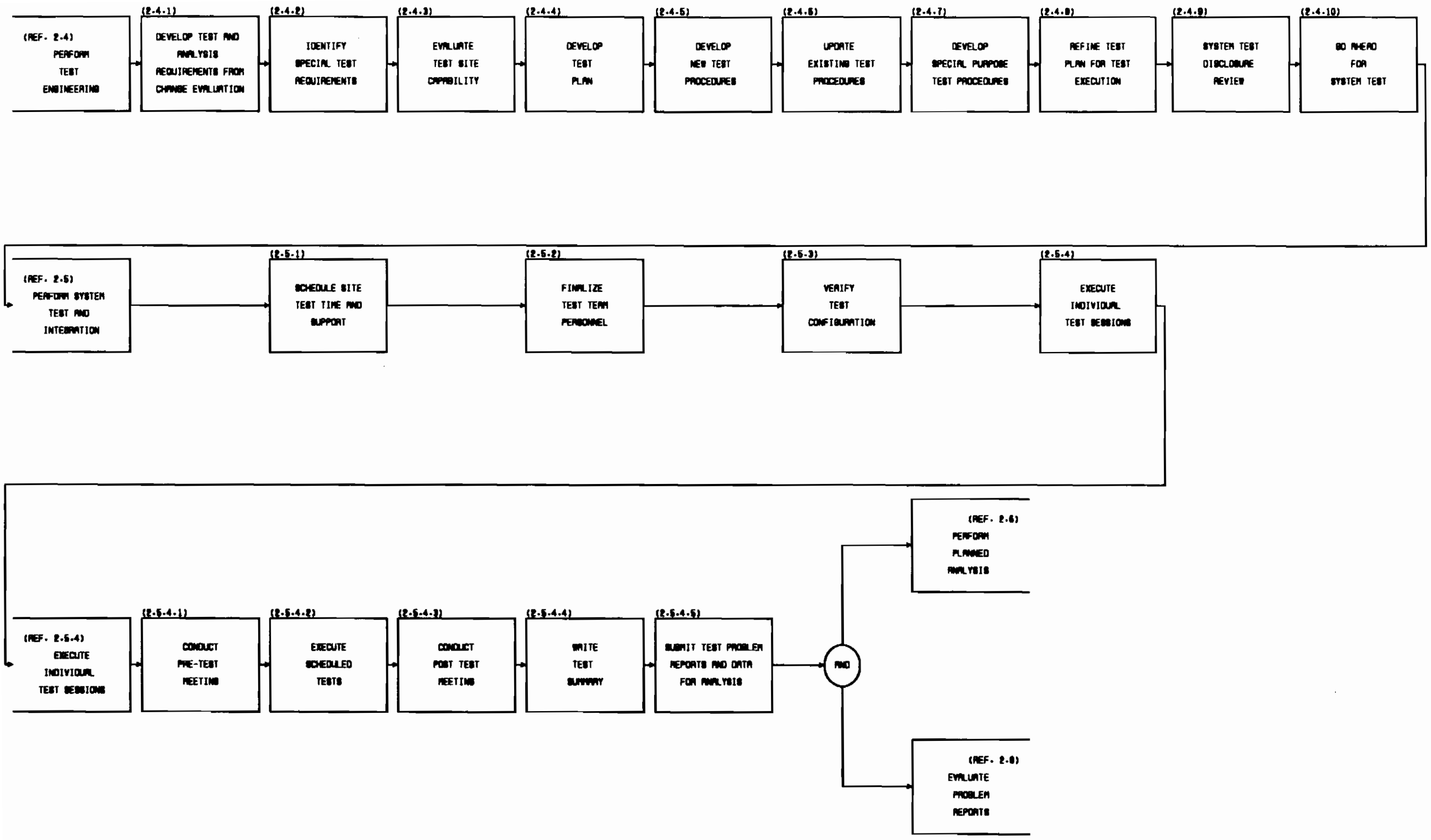

Figure 3. (continued)

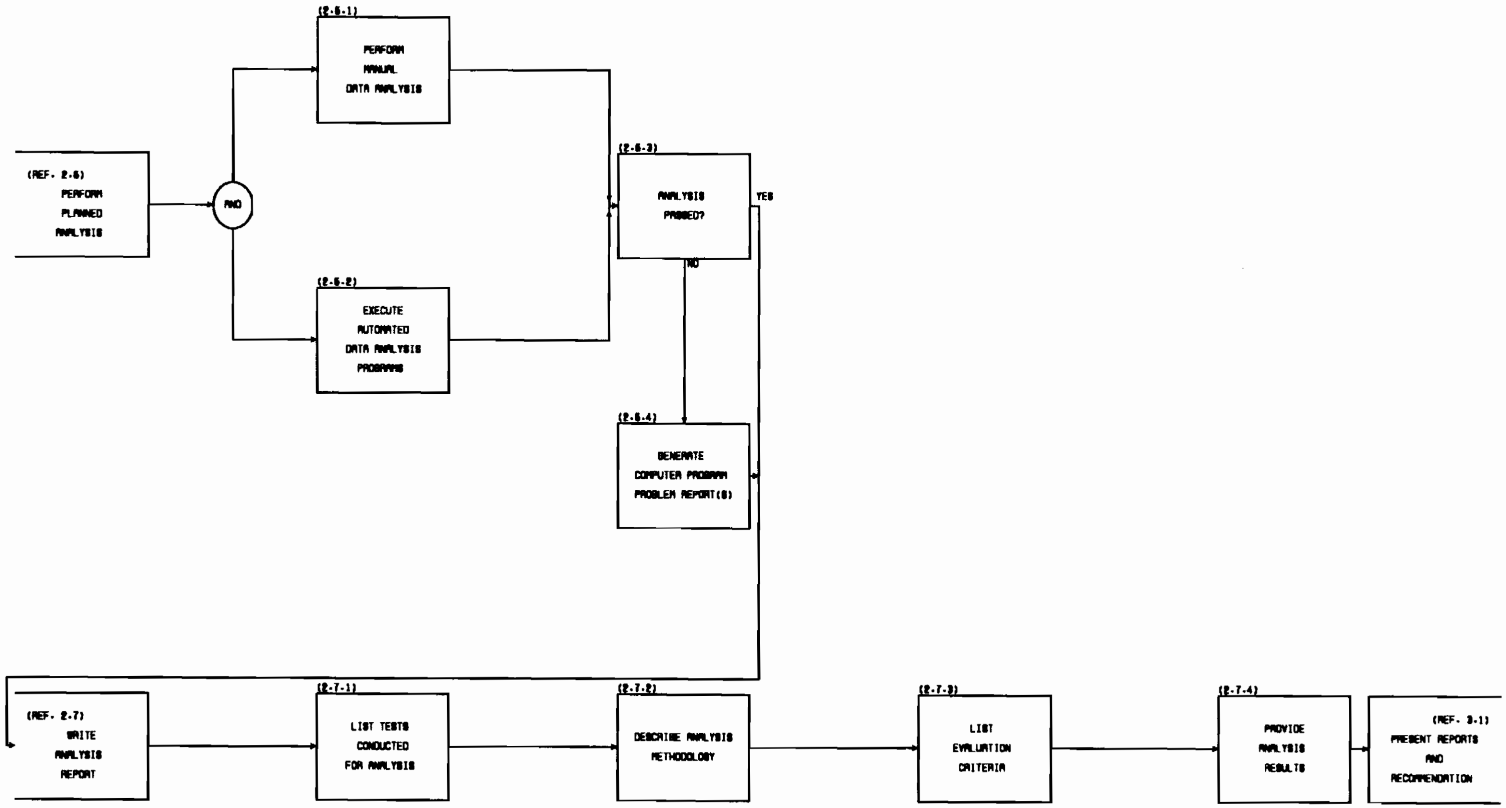

Figure 3. (continued)

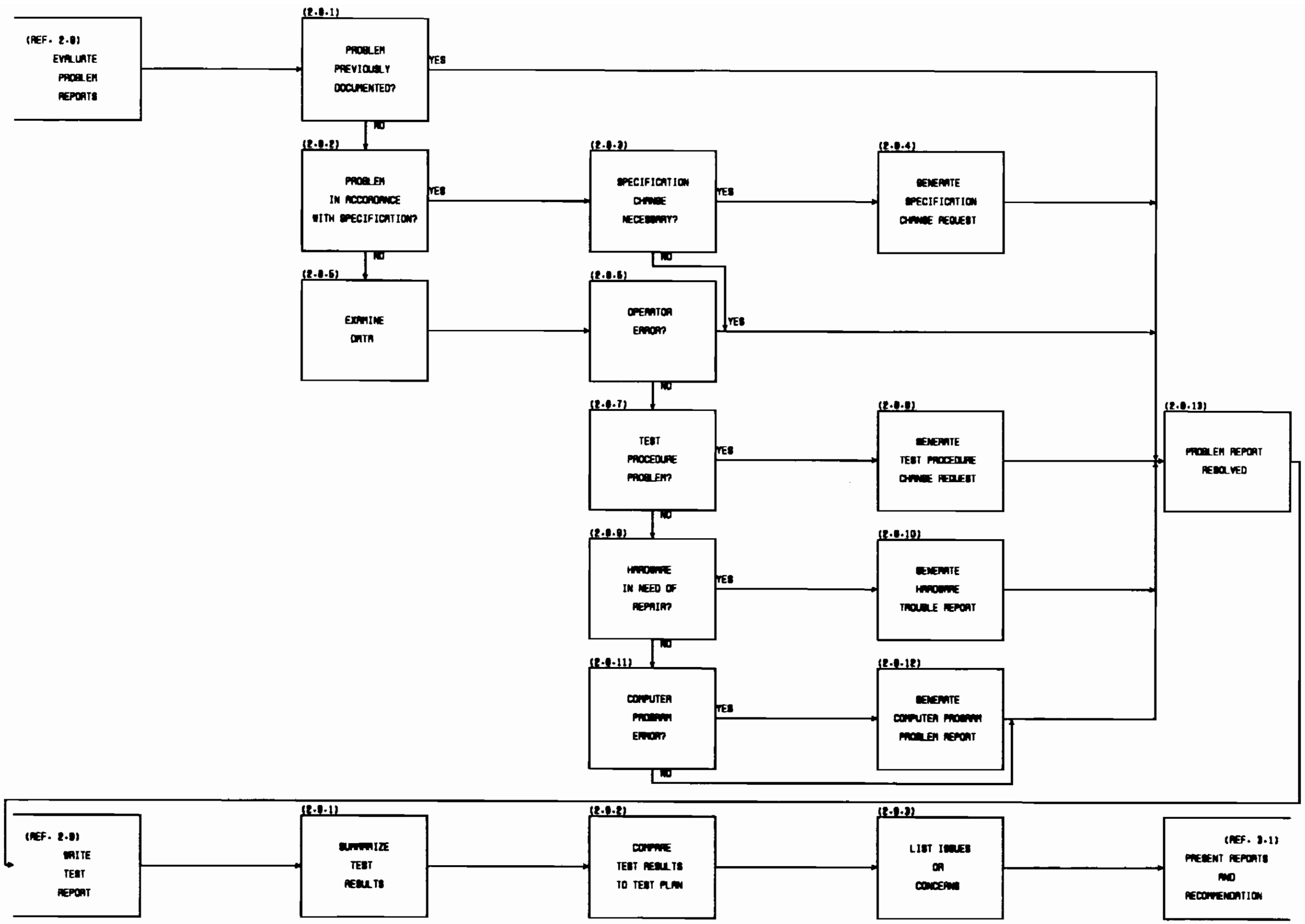

Figure 3. (continued)

mission areas are impacted? Is there new hardware involved? Is there operator impact? What impact is there on individual subsystems? The answers to these questions and others determine how the test engineering for the functional changes proceeds.

When Critical Design Review (CDR) has been completed (block 1.7), most of the core element changes have been identified. The approved specification changes and Computer Program Change Requests (CPCR) are received by the ST&I organization and evaluated for system impact (Reference block 2.2). A change is deemed to have system impact if: there is an IDS change associated with it, the change is coordinated between subsystems even though no IDS change is required, the change is to a mission capability, or there is possible high performance impact due to a combination of complexity, size, or affect on critical performance features. Once a change is deemed to have system impact it is assigned a priority of high, medium, or low. These priorities are considered during the test engineering phase as an indicator of test importance.

The integration of non-core element changes (Reference block 2.3) is often part of a major baseline upgrade. It is also a process that is frequently conducted independent of core baseline upgrades. In fact, test events driven solely by non-core changes may make up as much as 25 percent of all ST&I activity. The ST&I organization tries to link non-core integration efforts with major core upgrades to minimize testing. This is not always possible because of urgent non-core changes, the span of time between baseline upgrades, or the linking of non-core upgrades to hardware installation.

The process of non-core subsystem integration is continuous. ST&I personnel must constantly monitor the activities of all subsystem organizations. Integration begins with the coordination of schedules. ST&I personnel attend periodic subsystem planning meetings and maintain contact with the subsystem organizations. Compatible Combat System configurations are identified. Subsystem development schedules and ST&I schedules are coordinated and an ST&I window is agreed upon.

The ST&I organization then gathers subsystem change data. Core to non-core IDS changes are obtained. Non-core B-specification changes and CPCRs are obtained from the subsystem developers. If the subsystem or interface is new then new B-specifications or IDSs are obtained.

The subsystem changes are then evaluated for system impact (Reference block 2.3.3). The non-core subsystem change has system impact if: it is a new interface or subsystem, it is a new or changed functional capability, there is an IDS change associated with it, it is a coordinated change between subsystems, a mission capability is affected, or a high performance impact is possible.

If the subsystem changes do not have system impact then only regression testing is necessary. Regression testing is selective testing to verify that modifications have not caused unintended adverse side effects or to verify that a modified system still meets requirements.[9]

Test engineering for the non-core changes is conducted based on the results of the change evaluation. The test engineering philosophy is to establish a benchmark configuration and gradually add updated subsystems until the complete operational configuration is established. This minimizes the interaction of unknown entities and aids fault isolation.

Once all system level changes and subsystem level changes with system impact have been evaluated, the test engineering process is conducted (Reference block 2.4). Test and analysis requirements are developed to evaluate the changes. Any special test requirements, such as live aircraft or shipboard testing, are identified. LBTS availability and capability is also evaluated. With the test requirements and LBTS constraints in hand, the test plan is developed. New test procedures are developed if necessary, existing test procedures are updated or modified, and any special purpose test procedures are developed.

Once the test procedure package is complete and the time for ST&I draws near, the test plan is refined for execution. A System Test Disclosure Review (STDR) is held for the core LSEA organization. The ST&I test plan is presented and if adequate the go-ahead for system test is given. The next step is to perform the actual test and integration (Reference block 2.5).

With the go-ahead received, LBTS time and support is scheduled. Test team personnel assignments are finalized and a last-minute verification of the LBTS configuration is made. ST&I personnel are now ready for the actual execution of the individual test sessions (Reference block 2.5.4).

Test sessions are usually conducted in 4 to 6 hour blocks. A pre-test meeting is conducted by the test director one-half hour prior to each test block. All test personnel are assembled and briefed on the upcoming events. Test stations are assigned and responsibilities allocated.

The test director is in control during the execution of the tests. It is his or her responsibility to ensure: test objectives are met, all data necessary for analysis is gathered, problems are documented on problem reports, and a smooth flow of testing is maintained. "As-run" test procedures are maintained for each test. These procedures provide a record of the steps that passed, failed, or were not completed. A timeline of each test is also maintained. This aids analysis and re-creation of problems.

A post-test meeting is conducted immediately following each test block. All test personnel are assembled and debriefed. Problem reports are collected, analyzed for completeness, and assigned a priority by the test director of low, medium, or high. This priority is used by analysts to indicate the urgency of problem resolution.

It is the responsibility of the test director to write a test summary that assays the test event and highlights problems or concerns. All test problem reports and data are submitted for analysis. When tests are completed, the planned data analysis is performed (Reference block 2.6).

Some test procedures have sections that cannot be evaluated from visual results but are tailored for data extraction and analysis. The data analysis is performed either manually or through the use of automated data analysis programs. ST&I analysis personnel are responsible for developing many of the automated analysis programs. If an analysis test fails, the reasons for failure are investigated and if necessary a computer program problem report is generated.

The analysis report is written at the completion of analysis (Reference block 2.7). Presented in the report are: the list of tests conducted for analysis, the analysis methodology, the evaluation criteria, and the analysis results. The analysis report is then presented to the delivery review board.

All problem reports resulting from ST&I are analyzed and resolved (Reference block 2.8). Problem reports are assigned to specific personnel for analysis depending on the area or subsystem suspected in the report. Individual ST&I

personnel specialize in individual subsystems or warfare areas and are responsible for all problem reports in their area of specialty.

A test problem report in itself is not an indication that a computer program was at fault. The problem report is a result of an anomaly that an operator observed during testing. The anomaly may be the result of one of many factors. Each of these factors is considered when analyzing the problem report.

If a problem was previously documented and resolved then no further investigation is conducted. If the problem is found to be in accordance with specifications then the specification is evaluated. If the specification is thought to be in error or deficient then a specification change request is created. The data is examined. If the problem is attributed to operator error then the problem report is closed. If the test procedure is found to be in error then a test procedure change request is submitted. If the problem was caused by faulty hardware then a hardware trouble report is submitted to the LBTS. Finally, if a computer program is found to be in error then a computer program problem report is submitted to the appropriate change review board.

Once problem report analysis has been completed the ST&I organization prepares the test report (Reference block 2.9). The test report summarizes all testing and test results. Included are: test configurations, which tests passed, which tests failed, which tests were not completed, and the reasons for failed or incomplete tests.

The test results and activities completed are summarized and compared with the activities presented in the approved test plan. The satisfactory completion of those activities is a requirement for ST&I delivery recommendation. This includes the satisfactory results of tests for analysis. The programs must have properly implemented approved capabilities, must be stable in use, and must not regress from the previous versions.

In order to recommend delivery, all priority 1 and 2 computer program problem reports must be resolved. Department of Defense Standard 2167A classifies these priorities as follows:

> Priority 1. A software problem that does one of the following:
>
> (1) Prevents the accomplishment of an operational or mission essential capability specified by baselined requirements
>
> (2) Prevents the operator's accomplishment of an operational or mission essential capability
>
> (3) Jeopardizes personnel safety.
>
> Priority 2. A software problem that does one of the following:
>
> (1) Adversely affects the accomplishment of an operational or mission essential capability specified by baselined requirements so as to degrade performance and for which no alternative work-around solution is known

> (2) Adversely affects the operator's accomplishment of an operational or mission essential capability specified by baselined requirements so as to degrade performance and for which no alternative work-around solution is known.[10]

Any issues or concerns that resulted from the ST&I effort are also included in the test report.

Based on all of these criteria, the ST&I management makes a recommendation to the LSEA delivery review board (Reference block 3.1). The process of ST&I risk assessment is then complete. The review board will weigh the ST&I results and recommendation and decide if the Combat System is ready for delivery and installation. If the go-ahead is given, the system is installed on-board ship.

Key Problems Within the Process

Integration of non-core subsystems

The integration of non-core subsystems into the Combat System is a major part of ST&I. The Combat System complexity and number of configurations requires the process be executed as efficiently as possible. Since the LSEA is responsible for the development and maintenance of the core elements, it is natural that special attention is given to the integration of those elements. The concentration of effort on the integration of core elements may be done at the expense of non-core integration.

The test and integration of major core system upgrades is a well defined and well structured process. The assets of the entire LSEA are focused on the successful development, integration, and installation of the upgrade. When core subsystems are upgraded, the ST&I organization is an integral part of the process from conception to delivery. The organization has complete access to program managers, subsystem developers, documentation, change data, and information data bases. ST&I efforts related to the test and integration of core upgrades are scrutinized by the LSEA Quality Assurance organization.

The integration process for non-core subsystems is less structured. The successful integration of non-core subsystems depends almost entirely on the ST&I organization's effort. The LSEA Quality Assurance organization is concerned with the quality of core subsystems, it is not responsible for non-core subsystems. There may be little outside accountability and documentation required for the integration of non-core subsystems; the initiative of the ST&I organization is the driving force for successful integration.

The ST&I organization normally requires that subsystems be certified as operationally ready, by the developers, prior to ST&I. The ST&I organization may be forced into testing subsystems that have not yet been certified by the developer. This can occur if an integration test event has been scheduled but the subsystem development schedule has slipped. Programs that have not completed certification are likely to have problems that would have been caught in certification testing. The risk of testing uncertified subsystems is that the subsystem may not be in its final state. If it changes before delivery then it may be necessary to repeat integration tests with the new configuration.

The gathering of change data for non-core elements is more difficult and therefore may not be as complete as it is for core elements. Change data is not readily accessible to the ST&I organization. If it is accessible then it may not be in a format needed to make thorough impact assessment. The type of change information gathered from non-core program reviews and Computer Program Description Documents (CPDD) is usually not adequate to assess system impact. The information is often vague and lacks the necessary detail.

The evaluation of changes to non-core subsystems is not as thorough as it is for core subsystems. ST&I personnel may not be as familiar with the non-core subsystems as they are with the core subsystems; this is especially true if the subsystem or interface is new or if the subsystem has a major new functional capability. They may not have the information and assets available to make a complete change assessment. It is easy for ST&I personnel to communicate with programmers and subsystem developers within the LSEA, non-core developers are not as readily accessible.

The ST&I organization uses automated data collection and analysis programs extensively in the analysis of core element upgrades. Many of the analysis programs are developed by, or in conjunction with, the individual subsystem developers for use in both element level analysis and ST&I. This is done under the umbrella of the core LSEA organization. The ST&I organization may not have many automated programs for analyzing data gathered from the test and integration of non-core subsystems.

Management commitment to non-core integration is generally not as strong as it is for core elements. This is not surprising because the main function of the LSEA is the maintenance of the core Combat System computer programs. The core elements are managed within a single organization and are familiar. Even so, it is important to give the non-core elements the same level of attention necessary to ensure the integrity of the Combat System.

Test efficiency and effectiveness

For the reasons stated in Chapter 1, the ST&I organization must strive to improve the efficiency of test operations. LBTS time is at a premium. There is a large volume of tests to conduct within a limited schedule and personnel resources are scarce. These are all reasons to identify those activities that are inefficient and not value adding.

The LSE organization is often required to implement minor computer program changes to core subsystems to fix operational problems or to counter a new threat. For the same reasons, non-core subsystems may also implement emergency updates. When any subsystem changes, ST&I must be conducted to assess the impact on the Combat System. These minor test events may be conducted without any formal test plan or procedures. This ad hoc approach can result in the execution of tests without clear objectives and a well defined process for reaching those objectives. This type of testing lacks the efficiency of well planned tests and is ultimately less effective.

"Freeplay" testing is testing that is conducted without rigid test procedures. Freeplay testing may be used effectively to supplement structured system level testing. During system level freeplay testing, a proper environment is created and operators react to emergent situations. This is similar to the way the system would be used operationally. This type of system level freeplay testing is extremely dependent on the skills of the operators. The operators must have thorough knowledge of the system to create realistic situations. The operators must also recognize when a malfunction occurs and they must be able to re-trace the actions that created the malfunction.

Freeplay tests, in which operators exercise the system at their discretion, are in a large part ineffective for specific integration testing. The goal of these freeplay tests may be to "exercise" an interface or to have other subsystems "on-line" while conducting a primary test on different subsystems. The objectives of these tests are vague and operators often lack clear direction. Key requirements may be overlooked. If problems are found during freeplay testing they may be difficult to recreate because there is no record of the actions taken that led to the problem.

All freeplay testing is not ineffective. It is often used very effectively, with sound engineering judgement, to uncover problems that would not otherwise be detected. While freeplay testing is useful as a supplement to structured testing, the use of freeplay testing, as a method for testing specific system requirements or modifications, is not.

Inefficient test engineering fails to optimize the use of test operations personnel. Test procedure design may not take into account the number of personnel necessary to conduct the test and the operator loading during the test. A test procedure may require a test specialist who conducts only a few operations during the course of a test event and is idle the remainder of the time. Lack of test personnel cross training also fails to optimize personnel time. Poor test engineering and the lack of cross training results in the waste of scarce personnel resources.

The quality of test procedures directly impacts the effectiveness and efficiency of ST&I. A poor test procedure means that test requirements may be omitted or time may be wasted testing the wrong things. Procedures must be reviewed and updated continuously to ensure adequate coverage of system requirements. If procedures are not structured properly then tests do not flow in an efficient manner and test time is spent deciphering the test objectives. A poorly structured test procedure may not provide the proper environment, data, etc.. Retesting may have to be conducted that could delay delivery.

A test procedure library evolves over the life of a Combat System. Procedures are developed when a new configuration or requirement warrants them. The complexity of the Combat System and the large volume of test procedures contributes to the duplication of testing. A new procedure may test a requirement that is already covered in a different test procedure. If a test plan calls for the execution of both procedures then time is wasted testing the same requirement twice.

Some test procedures lack the modularity necessary to shape a test event to meet the test requirements and varying test schedules. A test procedure selected because it tests a certain requirement may also contain many steps that are superfluous. If the procedures cannot be easily separated into parts that match test requirements then unnecessary tests are conducted and ST&I resources are wasted.

Poor LBTS maintenance can cause much down time during test events. ST&I is the only LBTS user that utilizes the entire Combat System extensively. Training organizations rely heavily on the use of simulators and subsystem developers use only a small portion of the system at any one time. Hardware problems are found when ST&I users bring up the full system and exercise equipment that is infrequently used. ST&I operations must cease and ST&I personnel are idle while hardware troubleshooting and repair takes place. Because there is a fixed amount of LBTS time allotted to ST&I, the time is taken directly from test operations time.

Documentation and test event traceability

The ST&I organization executes the same tasks over and over. The same Combat System change may be made to many different ship configurations. It is necessary to conduct the same types of tests for each configuration. A comprehensive summary of results from the integration of similar system configurations aids the planning, execution, and analysis efforts. Situations that occurred during earlier integration events may arise in later events. If a good test history is not kept the organization could end up duplicating work that has already been completed.

Many small test events are conducted, the results analyzed, and the programs approved for delivery, all in a short period of time. This can be a result of dynamic ship schedules, emergency program updates, and uncontrollable work loads. Situations like these often require a decision to deliver the programs before all data analysis can be completed.

While an effective test process may have been completed, the time may not be taken to summarize, in one comprehensive report: the test event, specific configurations used, procedures used, and problems encountered. Certainly, records are kept in the form of site run sheets, problem reports, data tape logs, and other individual data points. The key is to summarize an event in one complete record. Chances are the same programs and problems will be encountered in a similar configuration down the road. If a complete and easily accessible record of past activity does not exist then the same process must be duplicated. This includes: requirements analysis, test engineering, test operations, and problem report analysis.

Documentation is a necessary part of the ST&I process. The organization is responsible for writing: test plans, test procedures, test reports, and analysis reports. Many times guidelines are not used for the orderly development and content of these documents. The lack of guidelines may be the cause of poor plans, procedures, and reports. Guidelines provide answers to the questions: What does a good test plan contain? How should a test procedure be structured? What is sufficient in a test report to provide traceability? When are the documents required? If the people that produce the documents don't know the answers to these questions then the documentation will not be adequate.

The ST&I organization does not participate extensively in subsystem level test operations. If the ST&I organization is not aware of problems that occurred during subsystem development testing then they may duplicate problem analysis efforts that were already conducted. Since subsystem developers are required to test each change that goes into an upgrade they must also develop procedures for those changes. The same procedures used during element level testing may be applied during system level testing. The only difference may be that the element level test was conducted using simulators and the system level test is conducted using real hardware. By not thoroughly tracking element level testing the ST&I organization is depriving itself of a valuable source of information.

Lack of a quality improvement program

Like many organizations, the ST&I organization may not follow a quality improvement program. The only way to keep up with the demands of the ST&I process is through the implementation of a program for continuous quality improvement. If no process is in place for continuous improvement then any quality gains realized today are going to be overwhelmed by the demands of tomorrow. The lack of strategic quality goals and processes to reach those goals may mean the organization has no long term vision but is only concerned with fighting the fires of the present.

Historically, the metrics used to judge the performance of an ST&I organization are not effective. They don't provide the data necessary to know whether or not the ST&I process is improving. Typical measures include such things as number of tests completed and percent of individual test completion. For example, the ST&I organization may have only completed 10-percent of a particular test but 95-percent of the key system requirements were tested within that 10-percent. The 10-percent completed is not as important as the 95-percent effectiveness. The lack of good metrics means the organization cannot effectively focus improvement efforts because it does not know precisely where the process is deficient and to what degree the process is deficient.

CHAPTER 4

HISTORICAL EXAMPLES OF NAVAL COMBAT SYSTEM TEST AND INTEGRATION

Naval Combat System Test and Integration is by no means an exact science. It is a continuously evolving process with many of the improvements in ST&I stemming from lessons learned during past events. The practice of ST&I is relatively new. The integration of large-scale, modern-day Combat Systems dates only to the early 1970s. The task of ST&I was drastically underestimated during those early years and turned out to be a painful learning experience. It is important to remember the mistakes made so we do not repeat those mistakes in the future. This chapter is a study of important events in the history of Naval ST&I.

CGN-36 (CALIFORNIA)

The following is from a technical report prepared for the Commander, Operational Test and Evaluation Force (COMOPTEVFOR):

> The failure of the Integrated Combat System of the USS CALIFORNIA (CGN-36), during her Acceptance Trials on 3-4 [sic] January 1974, caused a comprehensive introspective review, by the Chief of Naval Material, of the Navy's management practices in the acquisition of ships with modern sophisticated Combat Systems.[11]

The investigation focused on development testing and initial operational testing and evaluation but many of the lessons learned apply to the LSE phase of ST&I. In the end, the problems with hardware, software, and testing, that resulted in the Acceptance Trials failure, were traced to the inadequate management of those areas.

CALIFORNIA was the first surface ship designed with digital computers as an integral part of its Combat System. As the COMOPTEVFOR Report states: "The Chief of Naval Material stated that, as early as June 1968, the Navy recognized that the Combat Systems of the CGN-36 class represented a quantum increase in complexity over previous designs."[12] The magnitude of the ST&I task was still grossly underestimated. Many of the ST&I practices used today are a result of lessons learned from CALIFORNIA.

The COMOPTEVFOR Report summarizes the events leading up to the acceptance, from the shipbuilder, of CALIFORNIA:

> The T&E [Test and Evaluation] program for the Integrated Combat System of the CALIFORNIA was not adequate. . . . PCO [Prospective Commanding Officer] reports to CNO [Chief of Naval Operations] gave a continuous record of late accomplishment of scheduled events, of test deficiencies, and of hardware and software problems. During the summer of 1973, an awareness of crisis began to be appreciated; and in August 1973, the Navy recognized a state of "extremis" which led to the decision to accept the ship in an incomplete status. With only 4 months remaining until scheduled delivery of the ship, there probably was no other reasonable alternative; however throughout these last months, it was the event schedule that was the driving function, not the readiness of the ship to accomplish a particular event.[13]

There was no overall Combat System Manager (CSM) responsible for the ST&I program. For the most part, testing was conducted only by the individual subsystem organizations. Since many of the subsystems themselves were new or incorporated significant changes, the individual program managers had their hands full with internal problems. There were small scale attempts to run subsystems together at various LBTS. The COMOPTEVFOR Report describes one of these representative events:

> At Mare Island, on a "not-to-interfere-with-training" basis, and using the NAVSEC NTDS Test Tape, the NTDS/TARTAR D programs were run together. Simulation of inputs and outputs was attempted in lieu of major equipments. These tests did indeed uncover problems in both programs. However, the tests conducted were single thread tests and the systems were never saturated or tested in an operational sense. Development problems in systems, subsystems, and computer programs continued, many of which still existed when the programs and the equipments were delivered aboard ship.[14]

It was not until the subsystems landed aboard ship that ST&I was conducted in any fully integrated manner. It is true that there was no dedicated LBTS for CALIFORNIA Combat System development and test but the paragraph above is evidence the sites available were not used to their fullest capability. Again, the lack of a single point of contact for overall coordination precluded a structured program for land-based ST&I.

Inadequate documentation contributed to the poor state of the Combat System. There was no system level performance specification governing the development of the Combat System or the interfaces between subsystems. This meant there was no specification tree from the top level with functional allocation to lower level subsystems. In turn, there was no orderly way to judge the performance of the Combat System. System level test procedures could not be traced to any system level requirements.

It is evident that Test and Evaluation was not considered a major factor in the acquisition of CALIFORNIA. It never had been a major part in the Navy acquisition process and its importance was not realized until it was too late. A Combat System Test Plan (CSTP) including an Integrated Test Package (ITP) was provided to the shipbuilder in June 1970 as a contract modification.[15] Its implementation was not mandatory and, because there were constant modifications to subsystems, the procedures soon became obsolete.

There was no control over the individual subsystem computer programs. The COMOPTEVFOR Report paints a picture of indiscriminate "code slinging" by the subsystem developers:

> Configuration control of software programs was virtually nonexistent. Program changes were made by PARM personnel both at the LBTS and aboard ship and there was no coordinating authority to establish what effect these changes would have on the integrated performance of the total system.[16]

This situation is not surprising given the lack of central program management and governing system performance specification. Subsystems were not required to be operationally approved before system integration. At least one subsystem, the MK-74 (TARTAR D), was to obtain approval for operational use via its introduction on CALIFORNIA.

Test schedules for the Combat System were overly optimistic. In many cases the time required was at least twice the time allotted. "Among the reasons for delays were problems with spare parts, equipment, systems, and personnel."[17] Because of schedule pressures, many of the tests were conducted knowing the ship was not ready. Subsystem tests were overlapped with system tests. This undermined any effort for sequential testing. Analysis of test data also suffered because of schedule compressions.

CGN-38 (VIRGINIA)

The contract for construction of CGN-38 (VIRGINIA) was awarded on 21 December 1971.[18] Construction began at the same time the problems with CALIFORNIA began to surface. CALIFORNIA was evidence that an organized, systematic approach to ST&I was necessary to ensure the operational readiness of a complex Navy Combat System. Many of the lessons learned from ST&I of CALIFORNIA were applied to improve ST&I during construction of VIRGINIA.

A significant step toward the improvement of ST&I in the acquisition of Combat Systems was the issuance of DOD

Directive 5000.3 in January 1973.[19] This Directive established guidelines for Test and Evaluation in the acquisition of major defense systems. While it was too late to be applied to CALIFORNIA, it was effective for construction of VIRGINIA. The COMOPTEVFOR Report describes how DOD Directive 5000.3 and OPNAVINST 3960.10, which implemented the Directive for the Navy, impacted the Test and Evaluation of VIRGINIA:

> Although OPNAVINST 3960.10 which implemented these new policies in detail was not promulgated until Oct 1975, the context and content were well known early in the year. The T&E program for VIRGINIA was modified to conform; and even though a Test and Evaluation Master Plan (TEMP) for the CGN-38 was not officially signed until after ship delivery, the plan was effective in context for all later phases of the T&E effort.[20]

One of the most serious problems with the construction of CALIFORNIA was the lack of a central CSM that controlled the configuration and integration of the Combat System. For VIRGINIA, a CSM was designated in the Ship Acquisition Project Management (SHAPM) office and became the controlling authority over all subsystem efforts. The CSM was personally responsible for ensuring proper integration took place. This central authority prevented the piecemeal approach to Combat System Integration taken during CALIFORNIA construction.

Configuration management was also improved for VIRGINIA. CNO designated the Fleet Combat Direction Systems Support Activity (FCDSSA) as the Software Configuration

Control Manager that reported directly to the SHAPM. The COMOPTEVFOR Report describes the FCDSSA responsibility:

> Individual subsystem computer programs were required to undergo certification process. Once certified, they were delivered to FCDSSA for configuration control. In essence, the software programs were frozen and no further changes took place without positive identification of their impact on the other programs.[21]

This was an important step in ensuring the integrity of the Combat System. Once subsystems were certified, changes could not be made without impact assessment and appropriate test procedure development or modification. Specific subsystems that were not Approved for Service Use (ASU) were identified as high risk items. The PARMs for those subsystems were required to submit a plan of action for getting their programs ASU.

Test procedures were much better for VIRGINIA then they had been for CALIFORNIA. The procedures were validated at a LBTS before being handed over to the shipbuilder. The shipbuilder made better use of the procedures because he had more confidence in their quality. The contract was written in such a way that the shipbuilder had a better understanding of the requirements he had to meet. The contract specified that the ship had to meet installation requirements and higher level ST&I would be conducted after ship delivery.

Utilization of the LBTS at Mare Island was improved for VIRGINIA, although a dedicated LBTS still did not exist.

This meant the first time the computer programs were married with actual hardware, in a full Combat System configuration, was when they landed aboard ship. The steps taken with VIRGINIA greatly reduced the risks associated with waiting to conduct full ST&I until after ship delivery, but a dedicated LBTS would have ensured the computer programs for the Combat System were operationally ready before they reached the ship.

Land Based Test Sites (LBTS)

Neither CALIFORNIA nor VIRGINIA had a complete Integrated LBTS facility to conduct full-scale Combat System Integration. Hardware testing was conducted at various contractor facilities. For VIRGINIA, a shore facility existed at Mare Island for the purpose of computer program integration, test procedure check out, and limited crew training. Other sites with limited integration capability were located at: the Naval Surface Weapons Center (NSWC), Naval Ocean Systems Center (NOSC), Raytheon, and FCDSSA.[22] These sites were basically for subsystem development and none of them provided a complete Combat System configuration. Complete Combat System testing was conducted aboard ship. Other ship classes, such as LHA and DD-963, had similar limited LBTS facilities.

FFG-7 was the first ship acquisition project to incorporate a complete, dedicated LBTS for Combat System

Test and Evaluation. The FFG-7 LBTS was located at Islip, NY. It was constructed to provide a realistic environment for system level production testing, operational test procedure check out, crew training, and change assessment for the FFG-7 class of ships. The first ship was to be delivered in June 1977.

The FFG-7 Combat System Land Based Test Site (CSLBTS) was unique because it provided a central location where all the major subsystems were brought together. Full scale, physical mock-ups of the Combat Information Center (CIC) and equipment rooms were included. These mock-ups contained actual hardware as well as simulators. Functional radar, capable of tracking live aircraft, was also provided. Actual sensors and simulators provided console operators with realistic scenarios to test system effectiveness.

The FFG-7 CSLBTS was a prototype of the actual shipboard Combat System. Equipment and computer programs could be tested in an integrated system environment before production. The environment was highly controllable and was void of the problems associated with a ship under construction. For the first time, the complete Combat System could be tested before shipboard installation.

AEGIS is the Navy's state-of-the-art Combat System, it also has state-of-the-art LBTS facilities. The lead-ship for the AEGIS class became operational in 1983. Its Combat System had the advantage of being developed and tested at a comprehensive LBTS. The facility was constructed on the same premise as the FFG-7 facility, but the AEGIS Combat System is an order of magnitude more complex than the FFG-7 Combat System.

The Combat System Engineering Development (CSED) site for AEGIS is located in Moorestown, NJ. Its purpose is to provide a realistic Combat System configuration for the development, test, and evaluation of new AEGIS programs. The site also provides a realistic environment for coordinated team training for new ship crews. The site Officer in Charge (OIC) ensures fidelity of the Combat System configuration. Just as the shipboard configuration for new construction ships is tightly controlled so is the CSED site configuration. The CSED site mimics the actual hardware, to the greatest extent possible, and software that is installed in new construction ships. The site minimizes the use of simulators and emphasizes the use of operational hardware and computer programs.

The AEGIS Production Test Center (PTC) is used for production testing of hardware before it is installed on new construction ships. Each ship's Combat System hardware set is assembled and tested, as a system, before being shipped for installation. This increases the chance that hardware components will be compatible once they are installed.

AEGIS is also unique in the fact that there is a separate LBTS dedicated to LSE of AEGIS ships. The site is the Aegis Combat System Center (ACSC) located at Wallops Island, VA. Its purpose is to support the LSE functions necessary to sustain the AEGIS Combat Systems and provide new and refresher training for ship crews. The site is similar to the CSED site but with the emphasis on the testing of system upgrades, equipment retrofit, and investigating operational problems.

Because of its ocean-side location, the site not only has the ability to utilize live aircraft, as the CSED site does, but also can communicate and conduct exercises with ships at sea. It also differs from the CSED site because it must support numerous Combat System configurations to support all deployed configurations. The CSED site need only support the configurations under development.

The Navy has come to realize the importance of ST&I conducted at a LBTS. It is a concept born out of necessity. The Combat Systems of today are too complex and costly to risk failure after systems are installed. A comprehensive LBTS provides extensive ST&I capability, in a controlled environment, without the risk and cost associated with at-sea testing. A LBTS does not eliminate the need for operational at-sea testing, but it does eliminate much of the risk associated with it.

CHAPTER 5

A PLAN FOR IMPROVEMENT

The quality of the Combat System is not directly affected by ST&I; it is indirectly affected by the information gained from ST&I. Quality is designed into a system, it cannot be tested into a system. What ST&I can do is provide assurance the Combat System is fully capable of performing its mission. If the ST&I process is improved then the risk associated with deploying a Combat System is reduced. The ST&I organization is also a valuable source of feedback that is necessary in improving the Combat System. Future Combat System designs benefit from the knowledge gained through a high-quality Test and Evaluation program. If ST&I can provide better information about the quality of the Combat System then LSEA management can focus improvement efforts more effectively.

The way to improve the quality of the ST&I process is through a structured and comprehensive approach. Total Quality Management (TQM) is gaining a great deal of visibility in DOD these days -- and with good reason. TQM is a step beyond the traditional quality concepts of: product inspection, statistical quality control, and quality assurance. TQM is a way of thinking. When properly implemented, it touches every person and every activity within an organization. The DOD TQM Guide outlines the concept of TQM:

> Total Quality Management (TQM) is both a philosophy and a set of guiding principles that represent the foundation of a continuously improving organization. TQM is the application of quantitative methods and human resources to improve the material and services supplied to an organization, and the degree to which the needs of the customer are met, now and in the future. TQM integrates fundamental management techniques, existing improvement efforts, and technical tools under a disciplined approach focused on continuous improvement.[23]

TQM not only addresses the quality of the organization but also seeks to improve the quality of inputs into the organization. In the case of ST&I, the upstream organizations (see figure 2) include subsystem development and test organizations, change control boards, and system designers. Inputs from those upstream organizations are the element-certified subsystems, change packages, and specifications. If the quality of the inputs can be improved then the ST&I organization can be more effective. Time is not wasted analyzing and resolving problems from poor quality subsystems. Test engineering is easier if change packages and specifications are provided that contain the type of information needed to assess system impact.

Quality is viewed from the customer's perspective in a TQM system. Everything the organization does is to improve the quality of the product from the customer's point-of-view. The ST&I organization has two principle customers.

The LSEA managers use test results to make decisions with respect to the Combat System. More importantly, the fleet sailors receive and operate the integrated Combat System. Satisfying those customers should be the focus of ST&I efforts. If the Combat System conforms to specifications but does not meet user needs then it is not a quality Combat System. If all the test objectives of the ST&I organization are met but the information needed for management to make a confident decision is not provided then the process has failed.

There are many different step-by-step processes, found in quality literature, for implementing a TQM program. All of them outline an iterative process of: strategic planning, goal setting, taking action, measurement, analysis, and improvement. The DOD TQM Guide provides a seven step TQM model that can be used as a guideline and adapted to individual organization needs.[24] We will use this model to develop a plan for improving the ST&I process. The plan will include suggestions for improving some of the problem areas described in Chapters 3 and 4. Figure 4 depicts the TQM model.

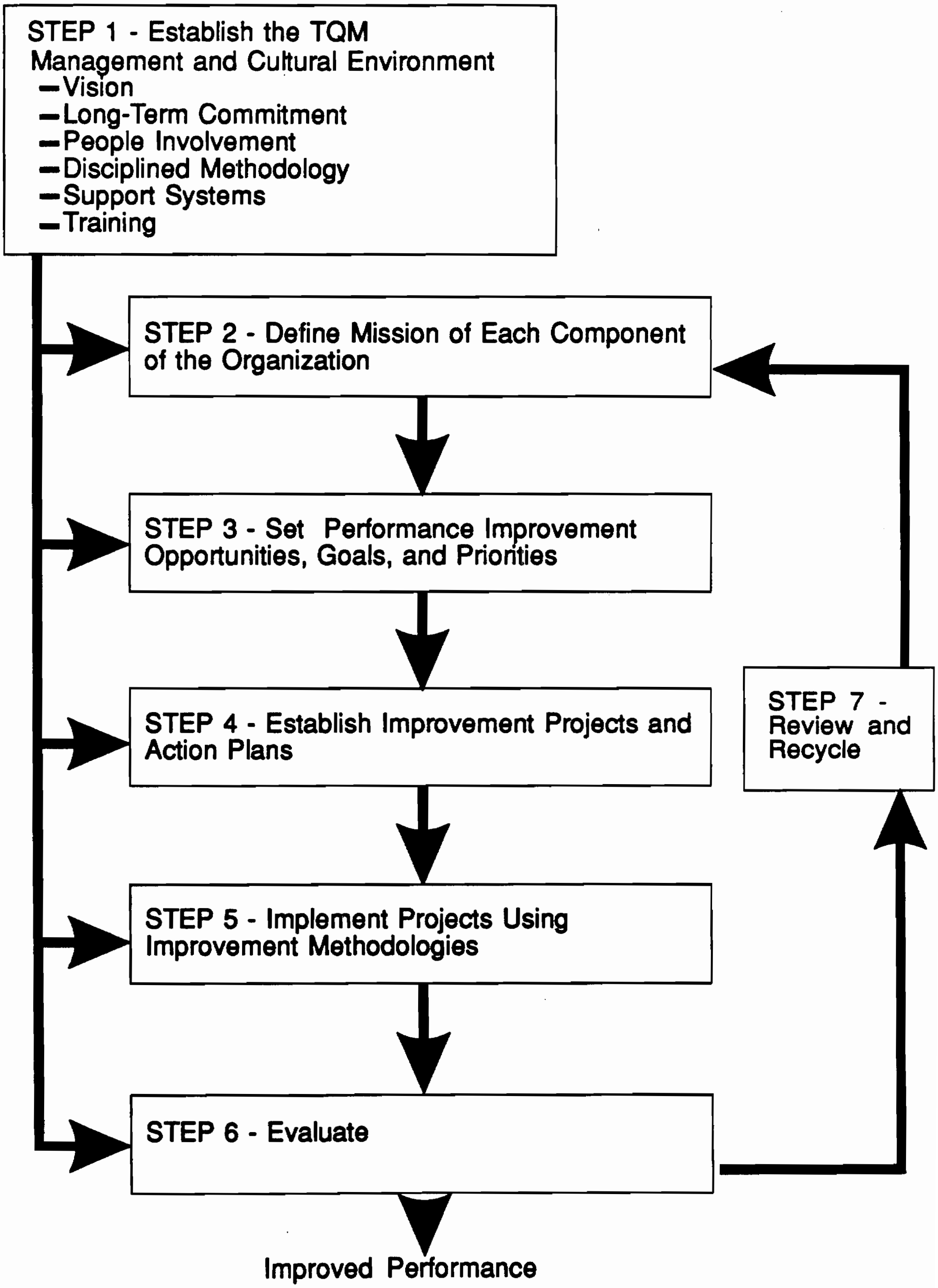

Figure 4. **Typical Total Quality Management Model [25]**

Step 1 - Establishing the Management and Cultural Environment

The first and foremost ingredient in a successful TQM program is commitment to quality by management. Management must recognize the need to improve and be willing to provide the leadership and support necessary for improvement. A clearly stated quality policy provides the vision of what the organization wants to be in the future. The policy is the foundation on which quality improvement programs are built. AT&T provides a good example of a quality policy:

> Quality excellence is the foundation for the management of our business and the keystone to our overarching goal of customer satisfaction. It is therefore, our policy to:
>
> - Consistently provide products and services that meet the quality expectations of our customers.
>
> - Actively pursue ever-improving quality through programs that enable each employee to do his or her job right the first time.[26]

Management must not only show commitment through policies but also through concrete actions. Employees lose faith when words are not backed by actions. Positive signs go a long way toward creating an attitude that the organization really is going to change for the better. ST&I managers must seize opportunities to show their commitment. Some examples of management commitment are: refusing to recommend programs that are not fit for use, not trading schedules for quality, and creating time for quality improvement projects.

There is naturally going to be resistance to change from employees. Management can help ease employee resistance to change by providing training opportunities, investing in tools for quality improvement, encouraging employee involvement, and providing recognition and/or rewards for meeting TQM goals. A step beyond employee involvement is giving employees the power to make decisions in their day-to-day work that improve quality. A free flow of information and ideas to and from management should be encouraged.

Discipline is essential in maintaining a TQM program. The organization must incorporate TQM into everyday activities. Initially processes must be standardized and then undergo continuous analysis to remove waste and simplify. New standards result from the analysis. The standards must be adhered to throughout the organization to create compatibility and eliminate ambiguity. The organization will continuously improve only if it has the discipline to continuously apply TQM methods.

Part of the environment needed for a TQM policy to be effective is an organizational support structure. Initially, it may be easiest to have the structure identical to the ST&I organizational chart, but providing for group interaction helps to focus on overall organizational goals rather than individual group goals. Upper management creates the environment of commitment, provides the vision and strategic goals, and serves as a steering group for middle managers. Middle managers set up improvement teams to tackle specific areas for improvement. They prioritize activities and provide the guidance to improve specific processes and functions. Improvement teams and individuals work on: skill development, specific process and function improvement tasks, and problem solving.

Training is an essential part of an effective TQM process. Since managers are responsible for motivating and educating employees, they must be trained in TQM concepts and methods. The degree of training an employee should receive depends on his or her organizational level and the nature of his or her tasks; therefore, the support structure should be established before large-scale training takes place. It is important to recognize that many of the tools used in industry will be difficult to apply to the ST&I organization and the organization must determine its goals and needs before selecting tools.

Step 2 - Defining the Mission of Each Component of the Organization

Chapter 3 describes the ST&I transformation process and all the activities that must be accomplished to transform individual subsystems into a synergistic Combat System. This step serves to allocate those activities to groups and individuals. The groups within the organization must understand, in detail, what their mission is with respect to the goals of the ST&I organization as a whole. Just as the organization has suppliers and customers, groups and individuals within the organization have suppliers and customers. The supplier or customer may be another group or a co-worker. For example, the test engineering group supplies procedures to the test operations group which in-turn supplies data to the analysis group.

The groups within the ST&I organization need to identify their customers and the present and future needs of their customers. The products and services that satisfy those needs should then be identified along with the processes used for their production. Communication within the organization is essential here. Each group plays an important role in the chain of events that lead to an integrated Combat System. If there is one weak link then there is a chain reaction that degrades the overall performance of the organization. The groups must work together to reach the goals of the entire organization.

The activities and outputs of each group should match its mission. If they do not match then processes need to be changed to reflect the mission. If an activity is not value-adding in terms of accomplishing the mission then it should be eliminated. If an activity is missing that is needed to accomplish the mission then it should be added. This exercise creates a sense of understanding of the purpose of each group and individual in the organization.

Every person will clearly know the role he or she plays in accomplishing the goals of the ST&I organization.

Step 3 - Setting Performance Improvement Opportunities, Goals, and Objectives

Step 1 creates a healthy environment for the TQM process. Step 2 identifies the role each group and individual will play to help the organization improve quality. With this foundation established the organization can set goals and objectives to improve quality. Upper management sets strategic goals for the organization that reflect the desire to satisfy customer requirements. Middle managers set goals for process improvement to achieve the strategic goals.

Chapter 3 identified some problem areas that should be addressed in order to improve the quality of the ST&I process. The key areas can be summarized in three strategic goals for improving the ST&I process. The goals are to: improve the integration of non-core subsystems, improve test efficiency and effectiveness, and improve documentation and traceability. A discussion of the three goals follows. Within these strategic goals, are more detailed goals and objectives that will help middle managers meet the strategic goals.

Improve the integration of non-core subsystems

This is an area that may be neglected in the LSE phase of ST&I for the reasons described in Chapter 3. It is important to continuously improve this part of ST&I because a major part of the Combat System comprises non-core elements and a large portion of ST&I time is dedicated to the integration of non-core elements. The integration of non-core subsystems is difficult because the subsystems are developed outside of the LSEA organization. This means extra effort must be made to gather information and assess the impact of non-core subsystems on the Combat System.

Step 2 of the TQM process will have identified those elements of the organization that play a role in the integration of non-core subsystems. It is up to the managers of those elements to set goals and objectives to improve the integration of non-core subsystems and create teams to accomplish those goals.

One of the goals should be to improve the communication with, and the coordination of, non-core activities. In TQM literature, this process is often referred to as vendor relations. TQM recognizes the fact that the quality of inputs from vendors has a direct impact on the quality of the product that is output from the transformation process. This is certainly the case with the integration of non-core subsystems. The quality of the subsystems and the information received from the subsystem developers directly impacts the quality of the integrated Combat System and the quality of the ST&I process. The experience of CGN-36 tells us that poor coordination of subsystem activities was a major contributing factor to the poor state of the Combat System.

Another goal is to improve the gathering of change information and the analysis of changes for non-core subsystems. This is a difficult process that must be continuously evaluated and improved. If changes to non-core subsystems are not thoroughly evaluated then the ST&I organization cannot completely assess the system impact of the changes. This increases the risk of deploying the Combat System.

Improving the capability for analyzing data gathered from the integration of non-core subsystems is another important goal. The only way to truly evaluate adherence to certain system requirements is through the analysis of data generated by non-core subsystems. The use of automated data reduction and analysis programs gives the ST&I organization the capability to analyze the data. It also expedites the resolution of problem reports and increases the accuracy of problem report information conferred to subsystem developers.

Improve test efficiency and effectiveness

Test efficiency is determined by the resources expended to complete the test event. Efficiency is important because it reflects how well LBTS time, personnel hours, and budget are used. Because of tight schedules and limited resources, time is a valuable commodity for the ST&I organization. It should be a continuing goal of ST&I to optimize test resources.

Effectiveness reflects the degree to which the right things were completed. For ST&I, effectiveness is the degree to which all tests necessary to assess Combat System risk were completed. If every test objective needed to conduct risk assessment was completed then the test event was 100-percent effective. Effectiveness, unlike efficiency, is not concerned with how many resources were used but whether or not all the right activities were completed. That is why it is important to consider test effectiveness along with efficiency with the understanding that tradeoffs may have to be made between the two.

Test procedures are the principle input to the test operations function. The quality of the procedures affects how well requirements are tested. Test procedures should be structured to ensure the correct requirements are tested and that resources are optimized in the process. Improving the test procedures will have a direct impact on test efficiency and effectiveness.

One of the goals to improve test procedures should be to standardize test procedure format. Test procedures are developed by many different people within the organization using many different methods; this causes variation in the quality of the procedures. Standardization means the procedures can be evaluated for content; evaluation is a fundamental part of continuous improvement. Standardization also means that operators will become more efficient in executing the procedures.

Another factor that influences test efficiency is the state of readiness of the LBTS. As stated in Chapter 3, ST&I uses configurations that the other LBTS users do not. Down time from hardware problems and system configuration problems takes LBTS time away from ST&I and wastes personnel hours. The ST&I organization expects some of this down time and may plan a "cushion" in the schedule to account for it. Rather than accepting the down time, a proactive approach should be taken to prevent it.

Improve documentation and test event traceability

Improving the documentation created by the ST&I transformation process should be one of the strategic goals of the organization. Test plans provide the guidance necessary for the effective execution of a test event. A good test plan also provides a benchmark to develop future plans for similar system configurations. Test reports and analysis reports compose the record of events the ST&I organization has conducted. They are the basis for making recommendations with regard to the Combat System. Reports also provide the type of comprehensive history necessary to easily recount past events.

Creating comprehensive test plans and test reports for every test event should be an objective of the ST&I organization. Test reports should be made available to subsystem developers, both core and non-core, to aid them in improving the quality of their subsystems. The reports will help the developers to better understand the ST&I process. The free flow of information will also improve working relationships between ST&I and the developers.

A goal for improving traceability should be to track not only the ST&I process but also the element level test process. Knowing the extent of testing conducted at the subsystem level provides a starting point for developing system test plans and procedures. A complete test history, all the way back to subsystem development testing, provides management with a complete record necessary to make decisions. Knowing a requirement was thoroughly tested at the subsystem level, even if it was omitted from ST&I, can provide the confidence necessary to recommend a computer program for installation. The give-and-take of ST&I providing test reports and a subsystem developer providing subsystem test information benefits both organizations and ultimately helps to improve the quality of the Combat System.

Step 4 - Establishing Improvement Projects and Action Plans

The greatest strategic resource the ST&I organization has is its people. The range of technical expertise within the ST&I organization is unequaled in any other LSEA organization. Because the ST&I organization is responsible for the integration of the entire Combat System, it has people with expertise in many different aspects of the Combat System. ST&I personnel also have a wide range of educational and practical experience, including: applied sciences, engineering, operational experience, and specific subsystem experience. The wide range of technical expertise means that given almost any task, some person or group of people can be found that is capable of completing that task.

In order to achieve the improvement goals of the organization, projects for improvement must be identified and selected. With the goals of the organization established in Step 3, specific improvement projects must be identified to realize those goals. The following are selected projects that will help the ST&I organization to achieve the goals of: improving the integration of non-core subsystems, improving test efficiency and effectiveness, and improving documentation and test event traceability.

Improving the integration of non-core elements

System integration planning is something that must originate at the program office level. Non-core subsystem maintenance schedules are controlled by the individual program offices. The core Combat System LSE schedule is controlled by the program office for the core Combat System. These program offices must work together to synchronize schedules and system configurations. The ST&I organization should be a key participant in the decision making process since these schedules drive ST&I events. If effective forums for system integration planning do not exist then the ST&I organization, with the aid of LSEA management, should help to create them by presenting the core Combat System program office with the need for the forums and a detailed plan for creating the forums.

To improve relations with non-core elements, ST&I technical personnel should be encouraged to participate in element level test events. This participation has a number of positive benefits. It increases insight into the extent of testing the subsystems undergo prior to delivery to ST&I. It provides a valuable training opportunity for ST&I personnel to learn from subsystem experts. It creates an environment of rapport and teamwork, between the subsystem developers and ST&I personnel, that is healthy for the maintenance of the Combat System.

For these same reasons, subsystem personnel should be encouraged to participate in ST&I events. It shows them the benefits of ST&I and the differences between ST&I and subsystem level testing. It provides them with the opportunity to see the role their programs play with respect to the rest of the Combat System. Since subsystem level testing is often conducted using interface simulators, it also allows them to see how their programs react to actual hardware and interfacing programs.

To realize an improvement in the analysis of non-core subsystem changes, the ST&I organization needs to create a structured methodology for gathering and analyzing change information. The organization must be proactive in gathering change information rather than relying on the type of information that historically has been provided by subsystem developers. ST&I personnel must decide what type of change information is needed to completely assess system impact then create a standard method for collecting the information and a process for analyzing the information.

Creating a standard is important since different subsystem developers provide different forms of change information. Once a methodology has been established, ST&I personnel must then work with the individual subsystem developers to gather the information. This gathering of information and impact assessment should be a part of the integration process for every subsystem program update.

To improve non-core subsystem analysis the ST&I organization must improve the tools for analyzing data gathered from the integration of non-core subsystems. Automated analysis programs increase the ability to analyze system impact and also increase the efficiency of problem analysis. Many subsystem developers use automated tools for analyzing data gathered during subsystem development testing. The ST&I organization should assess the usefulness of these tools for ST&I. A process should be established for creating and maintaining automated analysis programs with applications unique to ST&I.

Improving test efficiency and effectiveness

One project to increase the quality of test procedures is to develop a set of test procedure development guidelines. These guidelines should create standards for: the content of the test procedure, the sources for test requirements, provisions for tracing requirements, the degree of detail necessary, the optimization of personnel, and the level of modularity necessary for flexible test engineering.

Once test procedure standards are in place, existing procedures can be evaluated. A complete review of all test procedures should be conducted to establish the degree of compliance with the standards. One way of establishing and maintaining test procedure quality is by creating test procedure "champions". A champion will be responsible for ensuring his or her procedures meet the established standards and maintain an acceptable level of quality. This makes a specific person accountable for procedure quality and instills pride of authorship.

Action must be taken to ensure the LBTS is ready prior to the start of a test event. The first time block of a test event is often plagued by LBTS hardware problems, incorrect system configurations, and other LBTS logistics problems. Rather than having an entire test team idle while problems are sorted out, the ST&I organization should take steps to ensure the system is ready prior to the start of the event.

A skeleton crew of a few people, sent to the LBTS ahead of the main test team, would help to ensure the system is completely ready for the main event. Their job would be to load the entire suite of Combat System hardware and computer programs, ensure all communications and support equipment is operational, and take steps to correct any deficiencies. LBTS maintenance and support personnel should be present to assist the operation. A standard checklist would facilitate the system checkout.

If hardware failures are a chronic problem then the ST&I organization should work with LBTS management to ensure an effective preventive maintenance program is in place. Accumulating hardware failure data and other LBTS support system failures would provide the data necessary to point out chronic problems to LBTS management.

Improving documentation and test event traceability

A set of detailed guidelines should be developed for test plans, test reports, and analysis reports. These guidelines should describe the document format and content required. The circumstances that warrant the documents should also be defined. An analysis report may not be required for smaller test events. While test plans and test reports should be required for every event, the level of detail will depend on the scope of the event. A two day test event should not require a twenty page test report. The level of effort required should be such that the goal of creating a test plan and test report for every event can be realized.

A process for tracking subsystem testing should be developed in order to provide a complete test history. This is a necessary step to meet the goal of traceability. The type of information and the format for the information should be standardized. Standardization will ensure the correct information is available when needed.

In order for the information created and gathered to be of use, it must be readily accessible. There are many ways to cross reference: test plans, test reports, analysis reports, problem reports, data tapes, etc. Examples of the type of information needed are: the test history of a specific subsystem, problem reports associated with a specific subsystem, problem reports dealing with the interface between two subsystems, whether or not a certain configuration was tested, or test procedures utilized to test a specific configuration. An information system should be developed to correlate the documentation and data created by the ST&I organization. This would make the information more useful and minimize duplication of effort.

Summary of specific goals and projects

Table 1 is a summary of suggested strategic goals and lower level goals from Step 3 and specific improvement projects from this section. The ideas presented here do not compose a "cookbook" for improvement. The goals and projects were selected because they have the potential for significantly improving the ST&I process. Much of the organization's success depends on management commitment to improvement, how the projects are actually implemented, and the continuous evaluation and improvement of processes.

Step 5 - Implementing Projects Using Improvement Methodologies

Process improvement is realized through taking action. Once a project is selected, problem solving teams must be formed to implement the project. The teams should be trained in problem solving techniques. Structured group techniques are helpful for identifying the nature of a problem and also aid to provide solutions to the problem. Managers should track the progress of the improvement

Table 1. Suggested Goals and Projects for Improvement

Suggested Goals and Projects for Improvement		
STRATEGIC GOALS	**LOWER LEVEL GOALS**	**SPECIFIC PROJECTS**
Improve the integration of non-core subsystems	- Improve communication and coordination of non-core activities - Improve the gathering and analysis of change information - Improve non-core subsystem data analysis	- Develop a plan for creating forums at the program office level - Implement personnel exchange with subsystem developers - Create a structured methodology for gathering and evaluating change information - Develop a process for creating and maintaining automated data reduction and analysis programs
Improve test efficiency and effectiveness	- Improve test procedures - Standardize test procedure format and development methods - Improve LBTS readiness	- Develop a set of guidelines for test procedure development - Conduct a complete test procedure evaluation - Establish a LBTS checkout process

Table 1. (continued)

Suggested Goals and Projects for Improvement		
STRATEGIC GOALS	**LOWER LEVEL GOALS**	**SPECIFIC PROJECTS**
Improve documentation and test event traceability	- Create test plans and test reports for every test event - Track the element level test process	- Develop a detailed set of guidelines for documentation development - Create a process for tracking subsystem development testing - Develop an information system for correlating test information

project and provide guidance when necessary.

With specific projects established, a structured improvement process must be followed. The first step that must be taken is to identify the nature of the specific process and its: suppliers, inputs, outputs, and customers. This step is similar to the approach taken in Figure 2 to describe the entire ST&I process. A functional flow diagram like the one in Figure 3 is an excellent way to create a detailed description of a process. The process analysis should identify: the process owner and members, the role of the process members, the supplier requirements, and the customer requirements. The current process is then evaluated.

Benchmarking is one way of evaluating the current process. Benchmarking looks at a process compared with the process of competitors or the process of similar organizations. This can help the ST&I organization see where its processes stand with respect to those of similar organizations. Managers of similar processes are often willing to trade information that is mutually beneficial. Organizations that compare process information can take the best aspects from the aggregate and apply them to their own processes. Benchmarking is an activity that should be easier in DOD than it is in private industry. Information from quality leaders in private industry may be a secret that is guarded from competitors. The exchange of information for improving quality is encouraged within DOD.

Work Flow Analysis (WFA) is another method that can be used to identify unnecessary tasks and streamline processes. WFA identifies the steps used to complete a task -- possibly taken from the functional flow diagram. A cross-functional team analyzes the process. Lengthy steps, choke points, duplicate steps, and unnecessary steps are identified. Solutions to problem areas are identified and action plans for improvement are implemented.

The Nominal Group Technique (NGT) is an effective method for using a group to solve a problems. Groups are normally limited to a size of five to fifteen people. The NGT method is used to generate many ideas in a short period of time. The ideas are listed and clarified. Discussions are limited to clarifying ideas. Individuals prioritize the list of ideas. Votes are tabulated and an action plan is developed based on the selected high priority items.

Measures of Effectiveness (MOE) must be developed to evaluate the current process and focus improvement efforts. The performance measures must assess conformance to customer requirements. The MOE should also take into account the quality of the inputs and outputs. Data must be collected to assess the current process performance. The data is used to identify chronic problem areas that significantly degrade performance.

The next step is to find the causes for chronic problems. A cause and effect analysis, WFA, NGT, or other method can be used to identify problems that contribute to poor quality. These problems should be the focus of improvement efforts. A Pareto chart is useful to set priorities for improvement projects. A Pareto chart associates costs with problems and identifies problems that are the most significant contributors to poor quality. Priorities are then assigned to eliminate problems with the highest associated cost. Cost does not necessarily mean dollars, any common measure can be used. Figure 5 is an example of a Pareto Chart.

Once improvement projects have been identified and priorities have been set, the projects must be set in motion. The organization may want to test improvement projects before full-scale implementation. When a process is functioning as desired and gains have been realized, the process must be standardized in order to hold gains. The cycle of assessing the process, analyzing problems, identifying improvement opportunities, and implementing projects, is then repeated periodically. This creates a step function of process improvement and stabilization.

Step 6 - Evaluating

Evaluation is a necessary function at all levels of the organization. Measurement and evaluation is essential for assessing the impact of improvement projects. Project

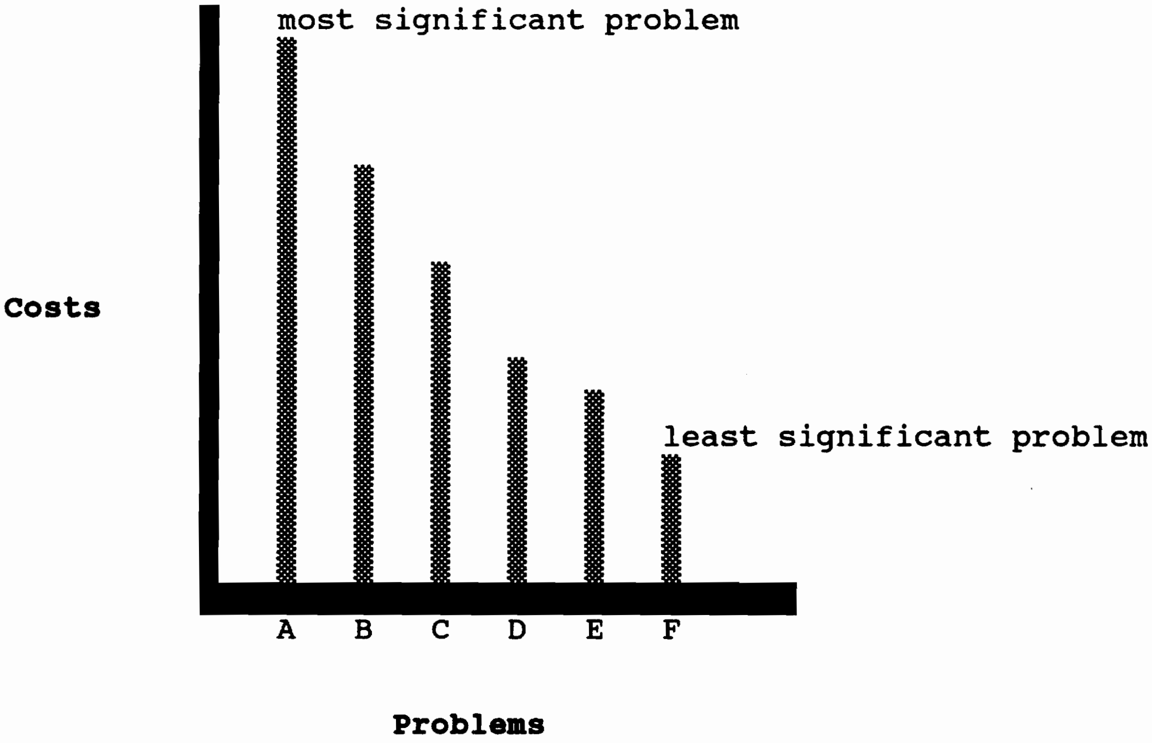

Figure 5. A Pareto Chart

measures are those measures developed by performance improvement teams in Step 5. Process measurement is used to track performance on a larger scale. Metrics that quantify the performance of the ST&I process will be discussed in detail in Chapter 6.

Behavioral change measures are used to assess the attitude of the people and environmental changes within the organization. Behavioral measures are less concrete than other types of measures but are no less important. The DOD TQM Guide lists the following factors to look for once a TQM program is implemented. There should be observable consistent evidence of:

1. Management support for continuous improvement
2. Trust between management and employees
3. Open communications without fear
4. Involvement of all employees
5. Teamwork
6. Supporting salary and reward system
7. Short-term issues do not overpower long-run issues
8. Process, rather than functional orientation
9. Knowledge and skills of TQM
10. Availability of time and resources for TQM
11. Employee support for TQM [27]

Continuous evidence of these factors shows management's commitment to the improvement of quality and employee confidence in the TQM process.

Step 7 - Reviewing and Recycling

TQM is a process for continuous improvement. The key word is continuous. TQM projects experience a slow start-up period, a period of growth in effectiveness, and then fading effectiveness. Once gains begin to fade, the cycle of analysis and improvement must be repeated. This forces TQM activities to a progressively higher level of effectiveness. For example: Initially, management may present ideas, ask for suggestions, and make a decision. This progresses to a more advanced stage where management presents problems, solicits suggestions, and makes a decision. A step beyond that is when management sets limits and asks a group to make a decision. Ideally, groups are empowered to independently work problems and make decisions.

This process of continuous review and improvement reinforces the idea that TQM is an everyday way of operating. The process is always improving. Instead of stagnation, projects are constantly undertaken to improve performance. The organization moves ahead instead of falling into a downward quality spiral. Every day, personnel must ask the questions: Are we performing as well as we can? If not, what can we do to improve?

CHAPTER 6

METRICS

Measurement is an essential component of process improvement. Metrics help the organization focus improvement efforts where they can be most effective. Metrics also provide the information necessary to evaluate current improvement efforts. For ST&I, measures of process effectiveness give managers the confidence to make decisions with regard to the Combat System. If measures are provided that show the ST&I process was effective, then decisions can be made with minimal risk.

The metrics presented here will center on those areas that are important for evaluating the ST&I process. Specifically, metrics will be provided that help to quantify: the effectiveness of test events, the efficiency of test events, the effectiveness of test procedures, the efficiency of test procedures, and the effectiveness of non-core subsystem integration.

These are certainly not the only measures of effectiveness for ST&I, but they are illustrative of the types of metrics that can provide valuable information for improvement. Remember, in order for metrics to effective, they must be easily understood, undisputed, and have a feasible means of gathering data. Most importantly, metrics must provide information that helps the organization meet its goal of continuous improvement.

Effectiveness of Test Events

The schedule of ST&I events is very dynamic. It is impacted by emergent program corrections, subsystem program delays, changing ship schedules, LBTS schedules, and a host of other factors. It is often necessary for the organization to make tradeoffs when planning a test event. It may not be possible to conduct all of the tests the organization would like, but a certain level of confidence must be achieved. Ideally, the organization would like to test every system level requirement, every subsystem interface requirement, and every possible combination of interface message traffic. Of course, one-hundred percent test coverage is not possible. The ST&I organization must establish test priorities and execute as many high-priority tests as possible within the allotted time. These priorities can provide guidance for test event engineering and can also provide a basis for evaluating test event effectiveness.

When planning a test event, the ST&I organization has a library of candidate test procedures. These test procedures, and subtests within the procedures, can be assigned priorities based on the test and integration task at hand. Test procedures can be divided into three priorities for planning and evaluation purposes. These levels distinguish between procedures that are critical, necessary, or supplemental.

The ST&I organization must develop clear guidelines for assigning these priorities. The following is an example of priority definitions. Critical procedures are those procedures that are essential for ensuring critical Combat System functions perform correctly. Necessary procedures are those procedures that are normally required for recommending delivery. Supplemental procedures are those procedures that are "nice to have," such as those that give the system more run-time but don't exercise any new requirements.

These priorities can be used at any level necessary to meet the organization's requirements. Priorities can be assigned to test procedures, subtests, specific changes, or specific requirements. The idea is for the organization to conduct tests in priority order as the schedule permits. Ideally, the organization will conduct all of the tests on its candidate list. In reality, some tests may be omitted.

For measurement purposes, the priorities are assigned numerical weights. The weights should be such that the failure to complete a critical event results in poor test event effectiveness and failure to complete a supplemental event has minor impact. This measurement model will assign the following priorities: critical (C) = 5, necessary (N) = 3, supplemental (S) = 1.

Table 2 and Table 3 are examples of how the test event effectiveness model is used. A value of 1 or 0 is assigned to a test depending on whether or not the test was completed. This Test Complete value is multiplied by the Weight of the test to get a Weighted Complete value for each test. The sum of Weighted Complete values is divided by the sum of test Weights to get a value for test effectiveness. In Example 1, 2 supplemental tests were not completed and the result was a value of 92.9% effectiveness. In Example 2, 2 critical tests were not completed and the result was a value of 64.3% effectiveness.

The model is impacted by the weights assigned to the individual priorities and also by the volume of testing. An ST&I organization must select appropriate weights to meet its needs and must also decide the level of effectiveness necessary to provide confidence in the test process.

Efficiency of Test Events

During a test event, many factors detract from the actual time spent executing the test plan. These factors include: system set-up, changing the system configuration, administrative delays, hardware down time, isolating emergent problems, etc. It is important for the ST&I organization to understand where the time goes during test

Table 2. Test Effectiveness Example 1

TEST	**PRIORITY C/N/S**	**WEIGHT**	**TEST COMPLETE**	**WEIGHTED COMPLETE**
SUBSYSTEM X/Y INTEGRATION				
-SUBTEST 1	C	5	1	5
-SUBTEST 2	C	5	1	5
-SUBTEST 3	S	1	0	0
SUBSYSTEM X REGRESSION				
-SUBTEST 1	C	5	1	5
-SUBTEST 2	C	5	1	5
-SUBTEST 3	N	3	1	3
-SUBTEST 4	N	3	1	3
-SUBTEST 5	S	1	0	0
COLUMN TOTALS		**28**		**26**
TEST EFFECTIVENESS = 26/28 = 92.9 %				

Table 3. Test Effectiveness Example 2

TEST	**PRIORITY C/N/S**	**WEIGHT**	**TEST COMPLETE**	**WEIGHTED COMPLETE**
SUBSYSTEM X/Y INTEGRATION				
-SUBTEST 1	C	5	0	0
-SUBTEST 2	C	5	1	5
-SUBTEST 3	S	1	1	1
SUBSYSTEM X REGRESSION				
-SUBTEST 1	C	5	0	0
-SUBTEST 2	C	5	1	5
-SUBTEST 3	N	3	1	3
-SUBTEST 4	N	3	1	3
-SUBTEST 5	S	1	1	1
COLUMN TOTALS		**28**		**18**
TEST EFFECTIVENESS = 18/28 = 64.3 %				

events. The information can be used to improve high cost areas. In the case of hardware and administrative delays, the information can be used to point out problem areas to the LBTS manager.

The ST&I organization keeps a time-line of the progression of a test event. This time-line contains useful information such as: the time when a particular test commenced, the time of hardware failures, and the time when problems were detected. This time-line process can be refined into a set of guidelines for tracking the progression of a test event. The time-line should include information that is important for improving the ST&I process.

Personnel training is important here. Personnel must understand what type of information is required and why the information is important. It is also essential that the information is not seen as a threat. The information should not be used to reprimand poor performance, but should be used in a positive way to improve processes.

Table 4 is a down-scaled example of a time-line breakdown for a four-day test event. An actual time-line would have more detail and would include additional information. The time-line shows start and stop times for important events. These events can be aggregated to provide a breakdown of how the LBTS was utilized. Table 4 shows that 11.25% of LBTS was dedicated to system set-up -- a

Table 4. Test Event Time-Line Example

TEST DATE	**TIME OF OCCURRENCE**	**DESCRIPTION**	**LENGTH OF OCCURRENCE**
6 May	1600-1645	System set-up	0.75 h
	1715-1745	Troubleshoot hardware	0.5 h
	2100	Test complete	
7 May	1600-1630	System set-up	0.5 h
	1800-1815	Locating special tape	0.25 h
	1815-1845	Reconfigure system	0.5 h
	2100	Test complete	
8 May	1600-1630	System set-up	0.5 h
	2100	Test complete	
9 May	1600-1630	System set-up	0.5 h
	1900-2000	Troubleshoot hardware	1.0 h
	2100	Test complete	
SUMMARY OF INFORMATION			
Category		Total Time	Percent Contribution
Actual test execution		15.50 h	77.50%
System set-up		2.25 h	11.25%
Hardware troubleshooting		1.50 h	7.50%
Administrative delays		0.75 h	3.75%
Scheduled test time		20.00 h	100.00%

candidate for improvement. While the information of one event may not be useful, the information gathered over an extended period of time can be used to isolate chronic problem areas. Each organization must tailor the type of information recorded to meet its needs.

Effectiveness of Test Procedures

The metric for evaluating the effectiveness of the test event does not take into account the quality of the test procedures. The ST&I organization may be very effective in executing high-priority tests, but if the tests exclude important requirements then Combat System risk assessment is not effective. A metric is needed that quantifies the coverage of specification requirements.

Just as test procedures can be assigned priorities, specification requirements can be assigned priorities for inclusion in test procedures. Some requirements are more important than others. While testing a specific interface message type may be important, merely changing a value within the message may be considered a supplemental requirement.

The ST&I organization must establish priorities for including system level requirements in test procedures. The organization must agree on the appropriate criteria used to assign requirement priorities. This task is best accomplished by a group. These priorities may be based on the specific requirement's contribution to the stated test procedure objectives. The priorities can be divided into the categories of High contribution (H), Medium contribution (M), or Low contribution (L). As in the test event effectiveness metric, weights can then be assigned to priorities. For example, assign the weights of High (H) = 10, Medium (M) = 5, and Low (L) = 1.

Evaluating requirements coverage requires a complete review of System Specifications and Interface Design Specifications. The specifications must be up-to-date and accurate. Each specification requirement is assigned a priority according to the guidelines. Once all of the pertinent requirements are evaluated, the test procedure is reviewed for requirements coverage. If a requirement is covered in the test procedure, it is assigned a value of 1. If the requirement is not covered, it is assigned a value of 0.

Table 5 is an example of how the metric is applied to a test procedure for the digital interface between two subsystems, Subsystem X and Subsystem Y. The IDS is the document that governs the interface between two subsystems. All possible message traffic between subsystems is contained in the IDS. Messages are divided into fields and the fields can take on several possible values. Each one of these messages, fields, and values are considered a requirement and can be assigned a priority. The Subsystem X/Y

Table 5. Test Procedure Effectiveness Example

TEST	**PRIORITY H/M/L**	**WEIGHT**	**REQT INCLUDED**	**WEIGHTED INCLUDED**
SUBSYSTEM X/Y INTEGRATION TEST				
-MT 101	H	10	1	10
-FIELD 1	H	10	1	10
-VALUE 1	M	5	1	5
-VALUE 2	L	1	0	0
-VALUE 3	L	1	0	0
-FIELD 2	M	5	1	5
-VALUE 1	L	1	1	1
-VALUE 2	L	1	0	0
-MT 102	H	10	1	10
-FIELD 1	M	5	1	5
-VALUE 1	M	5	1	5
-VALUE 2	M	5	1	5
-VALUE 3	L	1	0	0
-VALUE 4	L	1	0	0
-VALUE 5	M	5	1	5
-FIELD 2	M	5	1	5
-VALUE 1	L	1	1	1
-VALUE 2	L	1	0	0
-FIELD 3	L	1	0	0
-VALUE 1	L	1	0	0
-VALUE 2	L	1	0	0
-VALUE 3	L	1	0	0
-FIELD 4	H	10	1	10
-FIELD 5	M	5	1	5
COLUMN TOTALS		**92**		**82**

TEST PROCEDURE EFFECTIVENESS = 82/92 = <u>89.1 %</u>

Integration Test Procedure is then evaluated based on the how well the requirements are covered.

In reality, requirements are much more complicated than the example shows. Requirements are not always straightforward and may involve several conditions. Detailed requirements may be more applicable to element level testing than system level testing. These detailed requirements might even be assigned a weight of zero for ST&I purposes. Even though the process is cumbersome, it is a valuable exercise to evaluate how well the requirements are being covered. The ST&I organization will at least have the confidence that all requirements have been reviewed, are understood, and have been evaluated for system impact.

Efficiency of Test Procedures

One measure of test procedure efficiency is how well personnel resources are utilized. Personnel are the most important resource the ST&I organization has. Personnel not participating in test operations can be utilized to develop test procedures, conduct requirements analysis, etc.. If the number of personnel utilized to conduct test operations can be reduced, then the use of travel time, travel expense, and personnel hours, can be improved.

Test procedures are designed such that operators are assigned particular steps. If the degree of effort per step is consistent throughout the procedure, then the relative amount of effort required per operator can be calculated. This is a simple task. Total the number of steps for each operator. Divide the total for each operator by the total procedure steps. This is the percent contribution per operator.

If an operator provides little contribution to the procedure then his or her steps might be absorbed by another operator. If several test procedures are scheduled for a particular test event, comparing the operator contribution for each procedure can aid the test engineering process. If one test procedure requires little effort from a particular operator, but another procedure requires more effort, those two procedures are candidates for simultaneous execution.

Table 6 is an example of operator loading for two test procedures. In Test Procedure A, Operators 4 and 5 provide little contribution to the test; their actions are candidates to be absorbed by other operators, possibly operators 1 and 2. If both procedures are scheduled to be conducted, it may be possible to conduct them simultaneously since Operators 4 and 5 don't contribute much to Test Procedure A, but they are major players in Test Procedure B. Procedure interference and equipment usage are still factors for deciding which procedures to conduct simultaneously, but this metric can be used as one criteria. The metric can also be used to point out trends in operator usage. These trends suggest the areas for cross-training operators.

Table 6. Operator Loading Example

	TEST PROCEDURE A		**TEST PROCEDURE B**	
OPERATOR	**# ACTIONS**	**% CONTRIB.**	**# ACTIONS**	**% CONTRIB**
OPERATOR 1	21	11.4	0	0
OPERATOR 2	19	10.3	0	0
OPERATOR 3	45	24.5	0	0
OPERATOR 4	9	4.9	20	35.7
OPERATOR 5	3	1.6	28	50.0
OPERATOR 6	29	15.8	3	5.4
OPERATOR 7	26	14.1	5	8.9
OPERATOR 8	32	17.4	0	0
COLUMN TOTALS	184	100	56	100

Effectiveness of the Non-Core Subsystem Integration Process

There are many activities that must be conducted to complete the non-core subsystem integration process. Some of those activities are more critical to the integration process than others and each of the activities has a different measure of effectiveness. Quantifying the entire integration process requires the relative importance of individual activities be taken into account. It also requires the conversion of the measures for each individual activity to common units. This allows a Multi-Criteria measurement model to be developed. The Multi-Criteria measurement model allows dissimilar quantities to be aggregated into a common measure.

If the ST&I organization is following a plan for improvement, then the activities necessary for non-core subsystem integration will be well defined. Table 7 demonstrates how the Multi-Criteria metric is formed. Each subsystem integration activity can be assigned a relative degree of importance, or rating, with respect to the other activities. The most important activity is rated at 100 points. The other activities are rated relative to the most important activity. This process is best conducted using a group technique, such as NGT. Once the activities have been assigned ratings, the points are summed. For Table 7, the ratings total 660. The relative weight of each activity is found by dividing its rating by the total number of points.

Table 7. Multi-Criteria Evaluation for Non-Core Subsystem Integration Process

ACTIVITY PRIOR TO INSTALLATION	**RATING (1-100)**	**RELATIVE WEIGHT**	**DEGREE COMPLETED (SCORE)**	**WEIGHTED SCORE**	**POSSIBLE WEIGHTED SCORE**
1. ST&I personnel participated in subsystem testing	50	0.076	5	0.380	0.76
2. Subsystem test report received	60	0.091	0	0	0.91
3. ST&I gathered subsystem change information	100	0.151	7	1.057	1.51
4. ST&I evaluated subsystem change information	100	0.151	10	1.510	1.51
5. Subsystem element level test completed prior to ST&I	100	0.151	10	1.510	1.51
6. Subsystem personnel participated in ST&I	50	0.076	0	0	0.76
7. Analysis completed and problem reports resolved	80	0.122	10	1.220	1.22
8. Test report generated	70	0.106	10	1.060	1.06
9. Report distributed to subsystem personnel	50	0.076	10	0.760	0.76
TOTALS	**660**	**1.000**	---	**7.497**	**10.00**

Each subsystem integration effort requires that the individual activities be evaluated. Standards must be developed for evaluating these individual activities. The important point here is that whatever measure is used to evaluate the individual activities, that measure must be converted to a score on the scale of 1 to 10. This creates a common unit of measure for diverse activities.

Once the score has been decided for a particular activity, the score is multiplied by the weight of the activity. The weighted scores are summed and the result is the performance indicator for the integration event. This Multi-Criteria measure of effectiveness not only indicates the overall effectiveness of a particular integration event, but also points out weaknesses in the process. In the long run, scores from several events can be tabulated to isolate chronic problem areas.

This example must be tailored to the individual process. The rankings must be carefully set by the ST&I organization. These rankings determine how much of the integration effort should be devoted to each activity. For instance, activity 3 in the example, gathering subsystem change information, has a relative weight of 0.151. The simple implication is that 15.1% of the integration effort should be devoted to this activity.

It is also important that the organization develop standards for setting the scores of individual activities.

Setting these standards is another process that lends itself to a group technique. The group consensus is important because each member of the group, through his or her participation, is committing to living by the standards.

CHAPTER 7

CONCLUSIONS

Total Quality Management

TQM is a process that more and more organizations will find essential to keep up with increasing demands. If the demands of today are difficult to keep up with, what will organizations do to meet demands five years from now? TQM is not easy. Getting started is the hardest part. Today's ST&I organizations do not have enough time to complete their day-to-day projects, let alone add new projects to implement a TQM program. This lack of time is all the more reason to seek ways to improve. Eliminate processes that are non-value adding and improve efficiency. ST&I organizations constantly respond to high priority problems and apply resources to resolve those problems. Implementing processes for improvement to ensure sailors receive a safe, effective, high quality Combat System, must be recognized as a high priority task.

Quality is a spiraling effect. A downward quality spiral occurs when increasing test requirements result in: increasingly compressed schedules, less effective testing, less effective feedback to LSEA managers, and increased risk that poor quality programs may be approved for service use.

A TQM process can reverse this spiral. Improving processes means the ST&I organization can conduct more efficient ST&I. This means more time can be devoted to ensuring: the testing is effective, results are thoroughly analyzed, and effective feedback is provided to LSEA managers. This reduces the risk that poor quality programs will be approved for service use.

DOD is committed to TQM as a method for improving the quality of Combat Systems. Since DOD has made this commitment, some DOD installations have support offices dedicated to TQM. These offices provide training opportunities and support for those organizations that are new to TQM ideas and methods. The ideas presented in this report should be used as guidelines and examples. Individual organizations must use the resources available to learn TQM methods and develop TQM processes to meet their unique needs.

Future Research

The U.S. Navy acquisition process for Combat Systems has improved immensely since the days of CALIFORNIA and VIRGINIA. The acquisition process for new Combat Systems receives much visibility at high levels within the U.S. Government. Once those Combat Systems have been developed and deployed, they must be supported over their remaining life. More effort must be given to finding improved ways of conducting LSE for Combat Systems. Coordinating the activities that contribute to the LSE of a Combat System is particularly important.

There are many metrics that exist for assessing software quality during the development phase. Those metrics usually assess quality at lower code levels. Metrics are also needed to quantify quality at the system level. Metrics to assess Combat System quality would provide concrete measures for more complete risk assessment. System level quality is difficult to quantify because some measures of effectiveness, such as system useability, can be highly subjective. System level quality assessment is an important field of study that has widespread applications.

REFERENCES

1. Benjamin S. Blanchard and Wolter J. Fabrycky, *Systems Engineering and Analysis*, 2d ed., (Englewood Cliffs, N.J.: Prentice-Hall, 1990), 1.

2. William Hetzel, *The Complete Guide to Software Testing*, (Wellesley, Mass.: QED Information Sciences, Inc., 1984), 17.

3. Captain R. E. Adler and H. E. Mielo, "Combat Systems Integration: Engineering Discipline or Black Art," *NAVSEA Journal* (July 1976): 26.

4. Capers Jones, *Software Productivity Research*, (Dahlgren Va.: Naval Surface Warfare Center, N60921-91-DR6), videocassette.

5. J. M. Juran and Frank M. Gryna Jr., *Quality Planning and Analysis*, 2d ed., (New York: McGraw-Hill, 1980), 21.

6. Adler and Mielo, 24.

7. Jane's Information Group, Inc., *Jane's Fighting Ships, 1989-1990*, (Alexandria, Va.: Jane's Information Group, Inc., 1990).

8. Blanchard and Fabrycky, 632-640.

9. Hetzel, 17.

10. Department of Defense, *DOD-STD-2167A, Defense System Software Development*, (Washington, D.C.: U.S. Department of Defense, February 1988), 43.

11. Wheeler Industries Inc., "Utilization of Test and Evaluation Results in Weapons System Acquisition Process as Applied to Naval Surface Ship Systems," Final Technical Report W-78-N204-2 prepared for Commander, Operational Test and Evaluation Force, 30 December 1977, DTIC, AD-A048636, 4.

12. "Utilization of Test and Evaluation Results," 5.

13. "Utilization of Test and Evaluation Results," 39-40.

14. "Utilization of Test and Evaluation Results," 42.

15. "Utilization of Test and Evaluation Results," 43.

16. "Utilization of Test and Evaluation Results," 43.

17. "Utilization of Test and Evaluation Results," 42.

18. "Utilization of Test and Evaluation Results," D-1.

19. "Utilization of Test and Evaluation Results," 6.

20. "Utilization of Test and Evaluation Results," 6-7.

21. "Utilization of Test and Evaluation Results," 49.

22. Naval Sea Systems Command, "Total Ship Test Program for Ship Production (TSTP/SP); Report of the Ninth Conference," 18 June 1979, Naval Sea Systems Command, SEA-90E:MTR:jr.

23. Department of Defense, Total Quality Management Guide, [Final Draft], (Washington, D.C.: Office of the Deputy Assistant Secretary of Defense for Total Quality Management, 15 February 1990), 1.

24. Total Quality Management Guide, 15-47.

25. Total Quality Management Guide, 16.

26. AT&T Bell Laboratories, Quality by Design, Issue 1.1, (AT&T Bell Laboratories, 1978), iv.

27. Total Quality Management Guide, 45.

SELECTED BIBLIOGRAPHY

Adler, Captain R. E., and H. E. Mielo. "Combat Systems Integration: Engineering Discipline or Black Art." NAVSEA Journal, (July 1976), 24-26.

AT&T Bell Laboratories. Quality by Design. Issue 1.1. AT&T Bell Laboratories, 1978.

Blanchard, Benjamin S., and Wolter J. Fabrycky. Systems Engineering and Analysis. 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1990.

Defense Systems Management College, Test and Evaluation Management Guide. Fort Belvoir, Va: Defense Systems Management College, March 1988

Feeser, Henry Roger. "The Patrol Frigate Combat System Land-Based Test Site." Master's Thesis, Naval Postgraduate School, June 1974, DTIC, AD-781411.

Gallahue, James S.. "Combat System Test-Factory Through Shipboard." Naval Engineers Journal, (October 1980), 49-58.

Garvin, David A.. Managing Quality. New York: The Free Press, 1988.

Hetzel, William. The Complete Guide to Software Testing. Wellesley, Mass.: QED Information Sciences, Inc., 1984.

Jane's Information Group, Inc.. Jane's Fighting Ships, 1989-1990. Alexandria, Va.: Jane's Information Group, Inc., 1990.

Jones, Capers. Software Productivity Research, Dahlgren Va.: Naval Surface Warfare Center, N60921-91-DR6. Videocassette.

Juran, J. M., and Frank M. Gryna Jr.. Quality Planning and Analysis, 2d ed. New York: McGraw-Hill, 1980.

Li, Eldon Y.. "Software Testing In a System Development Process: A Life Cycle Perspective." Journal of Systems Management, (August 1990), 23-31.

Mull, Commander Charles L., USN. "Combat Weapon Systems-Problems and Promises." Naval Engineers Journal, (June 1970), 77-85.

Sink, D. Scott. Productivity Management: Planning, Measurement and Evaluation, Control and Improvement. New York: John Wiley & Sons, 1985.

U.S. Department of Defense, "DOD Directive 5000.2, Defense Acquisition Management Policies and Procedures, Part 8 - Test and Evaluation." Washington D.C.: U.S. Department of Defense, 23 February 1991.

U.S. Department of Defense. DOD-STD-2167A, Defense System Software Development. Washington, D.C.: U.S. Department of Defense, February 1988.

U.S. Department of Defense. Software Test and Evaluation Manual, Volume I, Guidelines for the Treatment of Software in Test and Evaluation Master Plans. October 1985. DTIC, AD-A165208.

U.S. Department of Defense. Software Test and Evaluation Manual, Volume II, Guidelines for Software Test and Evaluation in the Department of Defense. 25 February 1987. DTIC, AD-A163236.

U.S. Department of Defense. Software Test and Evaluation Manual, Volume III, Good Examples of Software Testing in the Department of Defense. 01 October 1986, DTIC, AD-A189425.

U.S. Department of Defense. Total Quality Management Guide. Final Draft. Washington, D.C.: Office of the Deputy Assistant Secretary of Defense for Total Quality Management, 15 February 1990.

U.S. Naval Sea Systems Command. "Total Ship Test Program for Ship Production (TSTP/SP); Report of the Ninth Conference." 18 June 1979, Naval Sea Systems Command, SEA-90E:MTR:jr.

Wheeler Industries, Inc.. "Utilization of Test and Evaluation Results in Weapons System Acquisition Process as Applied to Naval Surface Ship Systems." Final Technical Report W-78-N204-2 prepared for Commander, Operational Test and Evaluation Force, 30 December 1977, DTIC, AD-A048636.